MASTER'S DEGREE THESIS

MSc in Applied Mathematics

Title: Past, Present and Challenges in Yang-Mills Theory

Author: José Luis Tejedor

Advisor: Sebastià Xambó Descamps

Department: Departament de Matemàtica Aplicada II

Academic year: 2011-2012

Facultat de Matemàtiques i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

PAST, PRESENT AND CHALLENGES IN YANG-MILLS THEORY

José Luis Tejedor

June 11, 2012

Abstract

In electrodynamics the potentials are not uniquely defined. The electromagnetic field is unchanged by gauge transformations of the potential. The electromagnestism is a gauge theory with $U(1)$ as gauge group. C. N. Yang and R. Mills extended in 1954 the concept of gauge theory for abelian groups to non-abelian groups to provide an explanation for strong interactions. The idea of Yang-Mills was criticized by Pauli, as the quanta of the Yang-Mills field must be massless in order to maintain gauge invariance. The idea was set aside until 1960, when the concept of particles acquiring mass through symmetry breaking in massless theories was put forward. This prompted a significant restart of Yang-Mills theory studies that proved succesful in the formulation of both electroweak unification and quantum electrodynamics (QCD). The Standard Model combines the strong interaction with the unified electroweak interaction through the symmetry group $SU(3) \otimes SU(2) \otimes U(1)$. Modern gauge theory includes lattice gauge theory, supersymmetry, magnetic monopoles, supergravity, instantons, etc. Concerning the mathematics, the field of Yang-Mills theories was included in the Clay Mathematics Institute's list of "Millenium Prize Problems". This prize-problem focuses, especially, on a proof of the conjecture that the lowest excitations of a pure four-dimensional Yang-Mills theory (i.e. without matter fields) have a finite mass gap with regard to the vaccum state.

Contents

Chapter 1

Introduction

The quest for a unified theory of the physical world has been the goal of physics from the earliest days but especially since the beggining of the 20th century. Some well-known successes have been the unification of terrestrial and celestial gravity, the unification of electricity and magnetism, and the discovery that gravity was nothing but geometrical curvature. A full unification of gravitation, electromagnetism and the weak and strong forces, has still to be achieved. But some steps in that direction have already been taken and the most fundamental of these has been the discovery that all of the four fundamental interactions are governed by a single principle, namely, the *gauge principle*. The theory that embodies the gauge principle is known as gauge theory.

The discovery of the gauge principle as a fundamental principle of physics was a slow and tortuous process that took more than eighty years. It may be convenient to separate the discovery into three separate stages.

In the first stage it was shown, mainly by Hermann Weyl, that the traditional gauge invariance of electromagnetism was related to the coordinate invariance of gravitational theory and that both were related to the gauge invariance of differential geometry. Weyl was also the first to propose that gauge invariance be elevated from the rank of a symmetry to that of a fundamental principle. The second stage consisted in generalizing the gauge invariance used in electromagnetism to a form that could be used for the nuclear interactions. This stage culminated in the theory that is now known as Yang-Mills gauge theory. The third stage consisted of the gradual realization of the fact that, contrary to first apparences, the Yang-Mills theory, in a suitable modified form, was adequate for describing both of the nuclear inteactions. At this stage the difficulty was that the nature of the nuclear inteactions was masked by the low-energy phenomenology, and it required the introduction of a number of new and independent concepts (parity violation, spontaneus symmetry breakdown, color symmetry, asymptotic freedom, and so on) before their true character emerged. Even today, the evidence in the case of the strong interactions is only indirect.

The role of geometry in physics has always been central. But until the present century it was passive, providing only the stage on which the physics took place. The most profound entry of geometry into physics came with Einstein's theory of gravitation in 1916, which showed that gravity was nothing but the geometrical curvature of four-dimensional space. It was actually the theory of gravitation that opened the way for the developments in physics and mathematics that led to gauge theory. Although gauge theory is now universally accepted, its geometrical nature is not always as fully appreciated as the gravitational theory.

One of the important features of gauge theory is that its interactions are mediated by bosons

(particles with integer spin): gravitation by the graviton, electromagnetism by the photon, strong interaction by the gluons and the weak interaction by the vector fields $W^{\pm}$ and Z.

The goals of this work are:

1. to assess the importance of the gauge theory in modern physics by taking into account its evolution, with its successes and failures.

2. to show the characteristics and properties of quantum field theories. Modern gauge theories are quantum field theories and we present the two more known approaches: canonical quantization and path integral quantization. Furthermore, we consider the problem of renormalization (a program to eliminate the divergencies of a quantum field theory).

3. to describe the basic aspects of the Standard Model of the elementary particle physics and the quantum theories that synthesize our understanding of the fundamental interactions. These theories are quantum electrodynamics (QED), quantum chromodynamics (QCD) and the Weinberg-Salam-Glashow electroweak theory.

4. to acknowledge the existence of important unsolved problems such as the unification of gravity with the other three fundamental interactions, the unreasonable number of indeterminate parameters in the standard model or to prove the existence of a Yang-Mills theory on Minkowski 4-space.

We also consider some modern topics in gauge theories, such as lattice field theory, monopole solutions and supersymmetry.

Lattice field theory is the study of lattice models of quantum field theory, that is, quantum field on a spacetime that has been discretized onto the lattice. Although most lattice field theories are not exactly solvable, they have tremendous appeal because they can be studied by simulation on a computer. Lattice field theory keeps manifest gauge invariance, but sacrifies manifest Poincaré invariance –recovering it only after renormalization.

The quantum theory of the magnetic monopole started with a paper by the Paul Dirac in 1931. In this paper, Dirac showed that if a single magnetic monopole exists in the universe, then all electric charge in the universe must be quantized. In some gauge theories, genuine magnetic monopoles can be created as regular solutions of the field equations.

Traditional symmetries in physics are understood in the light of groups and their representations. The main groups involved are the Poincaré group (via its tensor representations) and those that encode internal symmetry. Supersymmetries, on the other hand, involve the spinor groups and their representations (spinor representations). Supersymmetry has become (along with superstring theory) the dominant framework for formulating physics beyond the Standard Model. But so far there is no experimental evidence neither of monopoles nor of supersymmetry.

Other topics such as instantons and superstrings are beyond of the scope of this work.

Finally, a comment on the main bibliography used for the preparation of this work. For chapter 2 (differential geometry and Lie groups), D. Bleeker [6] and F. W. Warner [12] and for chapter 3, L. O'Raifertaigh [9] and C. Misner, K. Thorne and J. A. Wheeler [10]. The book *Geometry and Quantum Field Theory* [13] was the main reference for quantum mechanics and F. Schwabl [11] for relativistic quantum mechanics. To introduce the canonical quantization of fields (chapter 5), L. H. Ryder [2], and S. Pokorski [3] for the path integral formulation of quantum field theory (chapter 6). To deal with the Standard Model, and particularly the electroweak model, the report of A. Pich [5] was very useful and also the book [16] by Cheng and Li. *Lectures on Supersymmetric Yang-Mills Theory and Integrable Systems* [14] was the basis for the study of

supersymmetry, and the books of M. Göckeler and T. Schücker [15] and L. H. Ryder [2], for magnetic monopoles. The reports of A. Jaffe [7, 8], and Arthur Jaffe and E. Witten [4] are the sources of chapter 9. The report of O. Garca-Prada [1] has been decisive in choosing the contents that I have developed in this work and otherwise most inspiring.

Chapter 2

Lie groups, connections and curvature

2.1 Lie groups and Lie algebras

2.1.1 Definitions and properties

Definition 2.1.1. Let G be an n-manifold and a group such that the groups operation $G \times G \to G$ given by $(g_1, g_2) \to g_1 g_2$ and the function $G \to G$ given by $g \to g^{-1}$ are C^∞ maps. Then G is called a ***Lie group***.

A *Lie subgroup* of a Lie group G is a subgroup $H \subset G$ which is also a submanifold of G.

Examples of Lie groups

- The $n \times n$ real invertible matrices form a group under multiplication, denoted by $GL(n, \mathbb{R})$ is a Lie group of dimension n^2, called the *general linear group*. The Lie subgroups of $GL(n, \mathbb{R})$ are called *matrix Lie groups*. Not only are they the most frequently encountered, but, because of the theorem of Ado and Iwasawa, practically anything which is true for matrix Lie groups has an analog for a general Lie groups. In fact, the groups that are relevant in the context of gauge theories consist of transformations represented by matrices, that can be parametrized in an analytic fashion in terms of a finite number of parameters. The number of independent parameters defines the *dimension* of the group.

- The rotation matrices 2×2, denoted by $SO(2, \mathbb{R})$. This group can be parametrized as follows:
$$SO(2, \mathbb{R}) = \left\{ \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} : \phi \in \mathbb{R}/2\pi\mathbb{Z} \right\}.$$

- *The Special Unitary Group.* The group $SU(n)$ is frequently used in elementary particle physics. Let $GL(n, \mathbb{C})$ be the space of all $n \times n$ matrices with complex entries. For $A \in GL(n, \mathbb{C})$, let $A^\dagger$ denote the conjugate of the transpose of A. Recall that the unitary group is $U(n) = \left\{ A \in GL(n, \mathbb{C}) | AA^\dagger = I \right\}$ and $SU(n) = \{A \in U(n) | \det A = 1\}$. If $t \to A(t)$ is a curve in $U(n)$ with $A(0) = I$, then (at $t = 0$) we have
$$0 = \frac{d}{dt}\left(A(t)A(t)^\dagger \right) = A'(0) + A'(0)^\dagger.$$
The dimension of $SU(n)$ is equal to $N^2 - 1$.

- The Lorentz group and the Poincaré group are the groups of linear and affine isometries of the Minkowski space. They are Lie groups of dimensions 6 and 10.

Definition 2.1.2. A *Lie group homomorphism* is a group homomorphism $\phi : H \to G$ which is also a smooth mapping of the underlying manifolds.

Definition 2.1.3. Let $\sigma \in G$. *Left translation* by σ is the diffeomorphism l_σ defined by $l_\sigma(\tau) = \sigma\tau$ for all $\tau \in G$. A vector field on G is called *left invariant* if for each $\sigma \in G$, X is l_σ-related to itself; that is, $dl_\sigma \circ X = X \circ l_\sigma$.

Definition 2.1.4. A *compact group* is a topological group whose topology is compact.

Basic examples of compact Lie groups include

- the circle group $\mathbf{T}$ and the torus group $\mathbf{T}^n$,
- the orthogonal group $O(n)$ and the special orthogonal group $SO(n)$,
- the unitary group $U(n)$ and the special unitary group $SU(n)$,

Every compact, connected, simple-connected Lie group K is a product of compact, connected, simple-connected simple Lie groups K_i.

Compact groups all carry a *Haar measure*, which will be invariant by both left and right translation (the modulus function must be a continuos homomorphism to the positive multiplicative reals, and so 1). In other words these groups are unomodular. Such a Haar measure is in many cases easy to compute; for example for orthogonal groups it was known to Hurwitz, and in the Lie group cases can always be given by an invariant differential form.

Definition 2.1.5. A *simple Lie group* is a connected non-abelian Lie group G which does not have non-trivial connected normal subgroups.

Definition 2.1.6. A ***Lie algebra*** is a vector space $\mathfrak{g}$ over some field F together with a binary operation $[\cdot, \cdot]$

$$[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g} \tag{2.1}$$

called the *Lie bracket*, which satisfies the following axioms:

- Bilinearity:

$$[ax + by, z] = a[x, z] + b[y, z], \quad [z, ax + by] = a[z, x] + b[z, y]$$

- Alternating on $\mathfrak{g}$: $[x, x] = 0$ for all $x \in \mathfrak{g}$
- The Jacobi identity: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$
  for all $x, y, z \in \mathfrak{g}$

The bilinearity and alternating propertialment imply anticommutativity, i.e: $[x, y] = -[y, x]$ for all elements $x, y \in \mathfrak{g}$.

Any vector space becomes a Lie algebra if all brackets are set equal to 0. Such a Lie algebra is called *abelian*.

The Lie algebra of the general linear group $GL(\mathbf{R})$ of invertible matrices is the vector space $M_n(\mathbf{R})$ of square matrices with the Lie bracket given by $[A, B] = AB - BA$.

Definition 2.1.7. Let $e_1, \ldots, e_n$ be a basis for the Lie algebra $\mathfrak{g}$ of G. The *estructure constants* $c_{ij}^k \in \mathbb{R}$ are defined by $[e_i, e_j] = \sum c_{ij}^k e_k$. Note that $[e_j, e_i] = -[e_i, e_j]$ implies $c_{ji}^k = -c_{ij}^k$. The Jacobi identity yields

$$\sum_m \left(c_{im}^h c_{jk}^m + c_{km}^h c_{ij}^m + c_{jm}^k c_{ki}^m \right) = 0 \quad \text{for all } h, i, j, k. \tag{2.2}$$

2.1.2 Exponential map

Definition 2.1.8. A homomorphism $\varphi : \mathbb{R} \to G$ is called a 1-parameter subgroup of G.

Definition 2.1.9. Let G be a Lie group, and $\mathfrak{g}$ be its Lie algebra. Let $X \in \mathfrak{g}$. Then

$$\lambda \frac{d}{dr} \to \lambda X$$

is a homomorphism of the Lie algebra of $\mathbb{R}$ into $\mathfrak{g}$. Since the real line is simply connected, there exists a unique 1-parameter subgroup $\exp_X :\to G$ such that

$$d \exp_X \left(\lambda \frac{d}{dr} \right) = \lambda X.$$

We define the *exponential map*

$$\exp{:}\mathfrak{g} \to G$$

by setting $\exp(X) = \exp_X(1)$.

Theorem 2.1.10. *Let X belong to the Lie algebra $\mathfrak{g}$ of the Lie group G. Then*

(a) $\exp(tX) = \exp_X(t)$ for each $t \in \mathbb{R}$.

(b) $\exp(t_1 + t_2)X = (\exp t_1 X)(\exp t_2 X)$ for all $t_1\,,\, t_2 \in \mathbb{R}$.

(c) $\exp(-tX) = (\exp tX)^{-1}$ for each $t \in \mathbb{R}$.

The exponential map from the Lie algebra $\mathfrak{gl}(n, \mathbb{C})$ of the general linear group $GL(n, \mathbb{C})$ to $GL(n, \mathbb{C})$ is defined by the usual power series:

$$\exp(A) = 1 + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

for matrices A. If G is any subgroup of $GL_n(\mathbf{R})$, then the exponential map takes the Lie algebrea of G into G, so we have an exponential map for all matrix groups.

2.1.3 Representations of Lie groups

Definition 2.1.11. A *representation* of G is a homomorphic map of elements of G onto matrices, $D(g)$ for $g \in G$. The representation matrices should satisfy $D(a)D(b) = D(c)$ if $ab = c$ for a, b, $c \in G$. The vector space v_j, on which representation matrices act, is called *representation space* such as $D(g)_{ij} v_j$ $(j = 1, \cdots, n)$. The dimension n of the vector space v_j $(j = 1, \cdots, n)$ is called *dimension* of the representation.

Definition 2.1.12. A subspace in the representation space is called *invariant subspace* if $D(g)_{ij}v_j$ for any vector v_j in the subspace and any element $g \in G$ also corresponds to a vector in the same subspace. If a representation has an invariant subspace, such a representation is called *reducible*. A representation is *irreducible* if it has no invariant subspace. In particular, a representation is called *completely reducible* if $D(g)$ for $g \in G$ are written as the following block diagonal form,

$$\begin{pmatrix} D_1(g) & 0 & & \\ 0 & D_2(g) & & \\ & & \ddots & \\ & & & D_r(g) \end{pmatrix} \tag{2.3}$$

where each $D_\alpha(g)$ for $\alpha = 1, \cdots, r$ is irreducible. This implies that a reducible representation $D(g)$ is the direct sum of $D_\alpha(g)$,

$$D(g) = \sum_{\alpha=1}^{r} \oplus D_\alpha(g)$$

2.1.4 Adjoint representation

We define the conjugation action as

$$c(g) : h \to ghg^{-1}$$

This action has fixed points, including the identity.

Definition 2.1.13. The differential of the conjugation action, evaluated at the identity, is called the adjoint action

$$Ad(g) = c_*(g) : T_eG \to T_eG$$

Identifying $\mathfrak{g}$ with T_eG and invoking the chain rule to show that

$$Ad(g_1) \circ Ad(g_2) = Ad(g_1 g_2)$$

this gives a homomorphism

$$Ad(g) : G \to GL(\mathfrak{g})$$

called the adjoint representation.

So, for any Lie group, we have a distinguised representation with dimension of the group, given by linear transformation on the Lie algebra. For the matrix group case, the adjoint representation is just the conjugation action on matrices

$$Ad(g)(y) = gYg^{-1} \tag{2.4}$$

since one can think of the Lie algebra in terms of matrices infinitesimally close to the unit matrix and carry over the conjugation action to them. Given any Lie group representation

$$\pi : G \to GL(V)$$

taking the differential gives a representation

$$d\pi : \mathfrak{g} \to End(V)$$

defined by

$$d\pi(X)v = \frac{d}{dt}(\pi(\exp(tX))v)|_{t=0} \tag{2.5}$$

for $v \in V$. Using the formula for the derivative of the differential of the exponential map, we find for the adjoint representation $Ad(g)$ that the associated Lie algebra is given by

$$ad(X)(Y) = \frac{d}{dt}(c(\exp(tX))_*(Y))_{t=0} = \frac{d}{dt}(Ad(\exp(tX))(Y))|_{t=0} = [X, Y] \tag{2.6}$$

For the special case of matrix groups we can check this easily since expanding the matrix exponential gives

$$e^{tX}Ye^{-tX} = Y + t[X, Y] + O(t^2) \tag{2.7}$$

So associated to *Ad(G)*, the adjoint representation of the Lie group G on $\mathfrak{g}$, taking the derivative we have $ad(\mathfrak{g})$, a Lie algebra representation of $\mathfrak{g}$ on itself

$$ad(\mathfrak{g}) \,:\, X \in \mathfrak{g} \to ad(X) = [X, \cdot] \in End(\mathfrak{g}) \tag{2.8}$$

As usual, the simplest example to keep in mind is $G = SU(n)$. In this case the Lie algebra $\mathfrak{su}(2)$ has a basis of skew-hermitician 2 by 2 matrices, these span the tangent space $\mathbb{R}^3$ to the group at the identity, which is that tangent space to S^3. The adjoint group action on this $\mathbb{R}^3$ is an action by orthogonal transformations in $SO(3)$. Using Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{2.9}$$

a standard base is $S_k = -i\sigma_k/2$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \tag{2.10}$$

where ϵ_{ijk} is a symbol antisymmetric in its indices and such that its 1 for ϵ_{123} and all cyclic permutations of the indices (123). Writing an element $X \in \mathfrak{su}(2)$ as

$$X = x_1\sigma_1 + x_2\sigma_2 + x_3\sigma_3$$

the adjoint group action on X by an element $g \in SU(2)$ is the map $X \to gXg^{-1}$ and this takes the vector $\mathbf{x} = (x_1, x_2, x_3)$ to a new vector $\mathbf{x}' = (x_1', x_2', x_3)$ where $\mathbf{x}' = A\mathbf{x}$ for some matrix $A \in SO(3)$.

The adjoint action of the Lie algebra on itself is given by the commutation relation for S_k

$$ad(S_k) : X \to [S_k, X]$$

One can work out what this means explicity in terms of matrices, for instance

$$ad(S_1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

2.1.5 Cartan subalgebra and roots

Let $\mathfrak{g}$ be a complex semisimple Lie algebra. Recall that $A \in \mathrm{End}(V)$ is called *semisimple* if it can be diagonalized.

Definition 2.1.14. $\mathfrak{h} \subset \mathfrak{g}$ is a *Cartan subalgebra* of $\mathfrak{g}$ if the following hold:

(a) $\mathfrak{h}$ is maximal abelian subalgebra.

(b) If $X \in \mathfrak{h}$ ad_X is semisimple.

Definition 2.1.15. Let $\mathfrak{g}$ be a complex semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$. Given a linear map $\alpha \in \mathfrak{h}^* = \mathrm{Hom}(\mathfrak{h}, \mathbb{C})$, define the *root space* of α as $\mathfrak{g}_\alpha = \{X \in \mathfrak{g} | [H, X] = \alpha(H)X$ for all $H \in \mathfrak{h}\}$. If $\mathfrak{g}_\alpha \neq 0$, α is called a root of $\mathfrak{g}$ with respect to $\mathfrak{h}$, or simply a root. Let $\Delta \subset \mathfrak{h}^*$ be the set of all non-zero roots of $\mathfrak{g}$.

The root spaces are one dimensional. The roots characterize the Lie algebra up isomorphism, and will encode a complete descrition in a diagram called the Dynkin diagram.

Theorem 2.1.16. *Let Δ be the set of roots with respect to a Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$. Then one has the following properties.*

(a) $\mathfrak{g} = \mathfrak{h} \oplus \sum_{\alpha \in \Delta} \mathfrak{g}_\alpha$.

(b) If $\alpha, \beta \in \Delta$, *then* $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}$.

(c) If $\alpha, \beta \in \Delta$, *and* $\alpha + \beta \neq 0$, *then* $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$.

(d) If $\alpha \in \Delta$, *then* $-\alpha \in \Delta$.

Definition 2.1.17. Let $\mathfrak{g}$ a Lie algebra and let $X, Y \in \mathfrak{g}$. The *Killing form* is defined by

$$(X, Y) = \mathrm{Tr}\, ad\, X\, ad\, Y \tag{2.11}$$

Remember that $\mathrm{ad}\, X$ is an operator which acts on elements of $\mathfrak{g}$ and maps them into new elements of $\mathfrak{g}$. Thus the indicated trace can be evaluated by first taking a basis for $\mathfrak{g}$, say $x_1, x_2, \cdots$. The we calculate for each x_j, the quantity $[X, [Y, x_j]]$ and express the result in terms of the x_i's. The coefficent of x_j is the contribution to the trace. The trace is independent of the choice of basis.

2.1.6 Root diagrams

The Killing form can be used to choose an orthonormal basis $h_1, \cdots, h_l$ of the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ and we extend to a basis $h_1, \cdots, h_l . g_1, g_2, g_{-2}, \cdots, g_{\frac{n-l}{2}}, g_{-\frac{n-l}{2}}$ of $\mathfrak{g}$ satisfying:

(1) $[h_i, g_j] = \alpha_i^j g_j \quad$ (no sum), $\quad \alpha_i^j \in \mathbb{R}$

(2) $[h_i, h_j] = 0$

(3) $[g_j, g_{-j}] = 0$

The basis elements g_j and g_{-j} are referred to as raising and lowering operators. Property (1) associates every g_j with l-tuple of real numbers $r^j = \langle \alpha_1^j, \cdots, \alpha_l^j \rangle$, called the roots of the algebra. Further, if r^j is a root, then so is $-r^j = r^{-j}$, and these are the only two real multiples of r^j which are roots. According to Property 2, each h_i is associated with the l-tuple $\langle 0, \cdots, 0 \rangle$. Because this association holds for every $h_i \in h$, these l-tuples are sometimes refered to as zero roots. For raising and lowering operators g_j and g_{-j}, Property 3 states that $r^j + r^{-j} = \langle 0, \cdots, 0 \rangle$.

Let Δ denote the collection of non-zeo roots. For roots r^i and $r^i \neq -r^i$, if there exist $r^k \in \Delta$ such that $r^i + r^j = r^k$, then the associated operators for r^i and r^j do not commute, that is

$[g_i, g_j] \neq 0$. In this case, $[g_i, g_j] = C_{ij}^k g_k$ (no sum), with $C_{ij}^k \in \mathbb{C}$, $C_{jk}^i \neq 0$. If $r^i + r^j \neq \Delta$, then $[g_i, g_j] = 0$.

When plotted in $\mathbb{R}^l$, the set of roots provide a geometric description of the algebra. Each root is associated with a vector in $\mathbb{R}^l$. We draw l zero vectors at the origin for the l zero roots corresponding to the basis $h_1, \cdots, h_l$ of the Cartan algebra. For the time being, we then plot each non-zero root $r^i = \langle \alpha_1^i, \cdots, \alpha_l^i \rangle$ as a vector extending from the origin to the point $\langle \alpha_1^i, \cdots, \alpha_l^i \rangle$. The terminal point of each rot vector is called *state*. As is commonly done, we use r^i to refer to both the root and the state. In addition, we allow translations of the root vectors to start at any state, and connect two states r^i and r^j by the root vector r^k when $r^k + r^i = r^j$ in the root system. The resulting diagram is called a root diagram.

EXAMPLE 1: $\mathfrak{su}(2)$

As an example, consider the algebra $\mathfrak{su}(2)$, which is classified as A_1. Using the Pauli matrices (2.9) we choose the basis $h_1 = \frac{1}{2}\sigma_3$ for the Cartan subalgebra h, and use $g_1 = \frac{1}{2}(\sigma_1 + i\sigma_2)$ and $g_{-1} = \frac{1}{2}(\sigma_1 - i\sigma_2)$ to extend this basis for all of $\mathfrak{su}(2)$. Then

(1) $[h_1, h_1] = 0$

(2) $[h_1, g_1] = 1g_1$

(3) $[h_1, g_{-1}] = -1g_{-1}$

(4) $[g_1, g_{-1}] = h_1$

By Properties 2 and 3, we associate the root vector $r^1 = \langle 1 \rangle$ with the raising operator g_1 and the root vector $r^{-1} = \langle -1 \rangle$ with the lowering operator g_{-1}. Using the zero root $\langle 0 \rangle$ associated with h_1, we plot the corresponding three points (1), (-1), and (0) for the states r^1, r^{-1}, and h_1. We then connect the states using the root vectors. Instead of displaying both root vectors r^1 and r^{-1} extending from origen, we have chosen to use only the root vector $r^{-l} = -r^l$, to connect the states (1) and (0) and (-1), respectively. The resulting root diagram is illustrated in Figure 1.

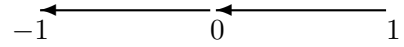


Figure 1. Root Diagram of $A_1 = \mathfrak{su}(2)$

EXAMPLE 2: $\mathfrak{su}(3)$

One representation of $SU(3)$ is the λ matrices of Gell-Mann:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \ \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \ , \ \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \ , \ \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \ , \ \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \ , \tag{2.12}$$

$$\lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \ , \ \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

The first three are just the Pauli matrices with an extra row and column added. It is convenient to work with the new basis.

$$\begin{aligned} T_\pm &= \frac{\lambda_1 \pm i\lambda_2}{2}, \qquad U_\pm = \frac{\lambda_6 \pm i\lambda_7}{2}, \qquad V_\pm = \frac{\lambda_4 \pm i\lambda_5}{2} \\ T_z &= \frac{\lambda_3}{2}, \qquad Y = \frac{\lambda_8}{\sqrt{3}}. \end{aligned} \tag{2.13}$$

where we have introduced the ladder operators. It is straightforward to compute all the commutation relations between the eight generators (we write only the non-zero commutators)

$$\begin{aligned} &[T_+, T_-] = 2T_z \qquad [T_+, U+] = V_+ \qquad [T_+, V-] = -U_- \\ &[T_z, T_\pm] = \pm T_\pm \qquad [T_z, U_\pm] = \mp U_\mp \qquad [T_z, V_\pm] = \pm V_\pm \\ &[Y, U_\pm] = \pm U_\pm \qquad [Y, V_\pm] = \mp V_\mp \qquad [U+, V_-] = T_- \\ &[U_+, U_\pm] = (3/2)Y - T_z \qquad [V_+, V_-] = (3/2)Y - T_z \end{aligned} \tag{2.14}$$

We can considerer these commutation relations between the eight generators to define the abstract Lie algebra of $SU(3)$. By examining the commutators carefully, we see that

- T_z and Y commute. This means the we can diagonalize both operators, and that the eigenstates are indexed by two numbers, called isospin and hypercharge.
- T_+ raises the eigenvalue t_z by one unit and T_- lowers it by one unit.
- Since $T_\pm$ commute with Y, they leave y the same.
- U_+ lowers the eigenvalue t_z by one-half unit and raises y by one unit.
- V_+ raises the eigenvalue t_z by one-half unit and raises y by one unit.

One basis of the Cartan subalgebra $\mathfrak{h}$ is $\{Y, T_z\}$ and we can calculate the roots easily. The root vector $\alpha_1 \equiv (1, 0)$ is associated to the raising operator T_+, $\alpha_2 \equiv (-1/2, 1)$ to the raising operator U_+, and $\alpha_3 \equiv (1/2, 1)$ to the raising operator V_+. And, of course, the opposite root vector will be associated to the respective lowering operators.

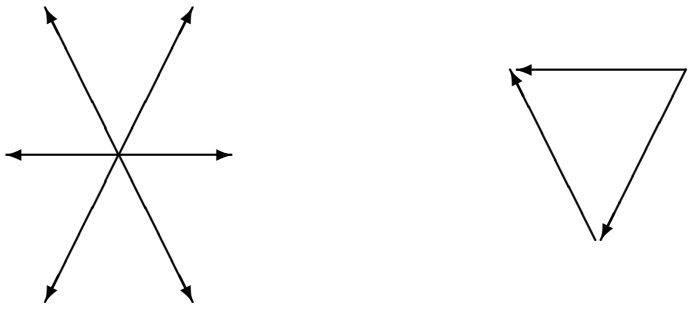


Figure 2. Root (left) and Minimal Weight Diagrams (right) of $A_2 = \mathfrak{su}(3)$

2.1.7 Weights and weight diagrams

An algebra $\mathfrak{g}$ can also be represented using a collection of $d \times d$ matrices, with d unrelated to the dimension of $\mathfrak{g}$. The matrices corresponding to the basis $h_1, \cdots, h_l$ of the Cartan subalgebra can be simultaneously diagonalized, providing d eigenvectors. Then, a list ω^m of l eigenvalues, called a *weight*, is associated with each eigenvector. Thus, the diagonalization process provides d weights for the algebra $\mathfrak{g}$. The roots of an n-dimensional algebra can be viewed as the non-zero weights of its $n \times n$ representation.

Weight diagrams are created in a manner comparable to root diagrams. First, each weight ω^i is plotted as a point in $\mathbb{R}^l$. Recalling the correspondence between a root r^i and the operator g^i, we draw the root r^k from the weight ω^i to the weight ω^j precisely when $r^k + \omega^i = \omega^j$, which at the algebra level occurs when the operator g^k raises (or lowers) the state ω^i to the state ω^j.

EXAMPLE: $\mathfrak{su}(3)$

The weight vectors of the considerated representation are

$$\phi^a = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \phi^b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \phi^c = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.15}$$

and their respective weight eigenvalue vector (left for T_z and right for Y) are

$$\omega^a = (0, -\frac{2}{3}), \qquad \omega^b = (\frac{1}{2}, \frac{1}{3}, 0), \qquad \omega^c = (-\frac{1}{2}, -\frac{1}{3}, 0), \tag{2.16}$$

The root and minimal non-trivial weight diagrams of the algebra $A_2 = \mathfrak{su}(3)$ are shown in Figure 2.

2.2 Lorentz group

The mathematical content of special relativity is that the Lorentz group, in fact the Poincaré group, (or at least the connected component of the identity in these groups) must act on the space of solutions to the equations of motion of particles and fields.

Definition 2.2.1. For $x = (x_0, \ldots, x_3)$ and $y = (y_0, \ldots, y_3) \in \mathbb{R}^4$, we define $\langle x, y \rangle = x_0 y_0 - x_1 y_1 - x_2 y_2 - x_3 y_3$.

According to special relativity, physical laws are unchanged by linear change of coordinates,

$$x'^\mu = \Lambda^\mu_{\ \nu} x^\nu + a^\mu \tag{2.17}$$

where Λ and a are real, provided it leave unchanged the invariant separation between points,

$$(x - y)_\mu (x - y)^\mu = \eta_{\mu\nu} (x - y)^\mu (x - y)^\nu \tag{2.18}$$

with $\eta_{\mu\nu}$ is a 4×4 diagonal matrix with entries $1, -1, -1, -1$ on the diagonal. This condition does not constrain a, since it cancels in the difference, but it imposes a constraint on Λ,

$$x_\mu x^\mu = x'_\mu x'^\mu = \eta_{\mu\nu} \Lambda^\mu_{\ \alpha} x^\alpha \Lambda^\nu_{\ \beta} x^\beta, \tag{2.19}$$

for all x^μ. A transformation of the form shown in (2.17) which satisfies (2.18) is called a *Poincaré transformation*. These transformations close and form a group, called the *Poincaré group*. The

subgroup where Λ is the identity matrix and a is arbitrary is a subgroup called the group of translations. We concentrate on the subgroup in which $a = 0$, which is called the *Lorentz group* and is denoted by $O(3,1)$. The minus sign in the metric create an important different between $O(4)$ and the Lorentz group: The invariant distance $s^2 = x_\mu x^\mu$ can be negative or positive, while the invariant distance of $O(4)$ was always positive. This means that the x_μ plane splits up into distinct regions that cannot be connected by a Lorentz transformation. If x and y are two position vectors, then these regions can be labeled by the value of the invariant distance s^2:

$$\begin{aligned}
(x-y)^2 = (x_\mu - y_\mu)\cdot(x^\mu - y^\mu) > 0: &\qquad \text{time-like} \\
(x-y)^2 = (x_\mu - y_\mu)\cdot(x^\mu - y^\mu) = 0: &\qquad \text{light-like} \\
(x-y)^2 = (x_\mu - y_\mu)\cdot(x^\mu - y^\mu) < 0: &\qquad \text{space-like}
\end{aligned} \tag{2.20}$$

The condition on $\Lambda^\mu_{\ \nu}$ to be a Lorentz transformation is,

$$\eta_{\mu\nu} x^\mu x^\nu = \eta_{\alpha\beta}\Lambda^\alpha_{\ \mu}\Lambda^\beta_{\ \nu} x^\mu x^\nu \tag{2.21}$$

for all x^μ. Since this must hold for all x^μ, we have

$$\eta_{\mu\nu} = \Lambda^\alpha_{\ \mu}\,\eta_{\alpha\beta}\,\Lambda^\beta_{\ \nu}, \tag{2.22}$$

or using matrix notation

$$\eta = \Lambda^T \eta \Lambda \tag{2.23}$$

The group of matrices Λ satisfying (2.23) is called $O(3,1)$ and is a Lie group. The Lorenz group has 4 disconnected pieces. To see this, take the determinant of (2.23):

$$\det\eta = \det\Lambda^T\eta\Lambda = \det\eta \times (\det\Lambda)^2 \tag{2.24}$$

Since η is non-singular, we can divide by $\det\eta$ and $\det\Lambda = \pm 1$. An element of $O(3,1)$ with $\det\Lambda = 1$ is called *proper*, and an element with $\det\Lambda = -1$ is called *improper*. Furthermore, if we write out $\mu = 0$ and $\nu = 0$ element of (2.23), it is

$$(\Lambda^0_{\ 0})^2 = 1 + \sum_{i=1}^{i=3} (\Lambda^i_{\ 0})^2 \geq 1, \tag{2.25}$$

so the square of the time-time component of any Λ must always be at least 1, and Λ^0_0 must be either ≥ 1 or ≤ 1. An alement of $O(3,1)$ with $\Lambda^0_0 \geq 1$ is called *orthochronous*, and an element with $\Lambda^0_0 \leq -1$ is called *non-orthochronous*. The *parity transformation* $(t,\mathbf{x}) \to (t,-\mathbf{x})$ is a canonical example of improper but orthochronous transformation. The *time reversal*s $(t,\mathbf{x}) \to (-t,\mathbf{x})$ is a canonical example of improper and non-orthochronous transformation. PT (parity and time reversal) is a example of proper and non-orthochronous transformation.

The four connected components of the Lorentz group L are

$$\begin{aligned}
L^\uparrow_+ &= \left\{\Lambda \in L |\ \det\Lambda = 1, \Lambda^0_{\ 0} \geq 1\right\}, \\
L^\downarrow_- &= \left\{\Lambda \in L |\ \det\Lambda = -1, \Lambda^0_{\ 0} \geq 1\right\}, \\
L^\uparrow_+ &= \left\{\Lambda \in L |\ \det\Lambda = 1, \Lambda^0_{\ 0} \leq -1\right\}, \\
L^\downarrow_- &= \left\{\Lambda \in L |\ \det\Lambda = -1, \Lambda^0_{\ 0} \leq -1\right\}.
\end{aligned} \tag{2.26}$$

2.2.1 Generators of the Lorentz group

A infinitesimal Lorentz transformation must be of the form

$$\Lambda^{\mu}_{\ \nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\ \nu}, \tag{2.27}$$

with $\omega^{\mu}_{\ \nu}$ a matrix of infinitesimal coeficients and δ^{μ}_{ν} is identity matrix. The condition on $\omega^{\mu}_{\ \nu}$ for $\Lambda^{\mu}_{\ \nu}$ to be a valid Lorentz transformation is found by inserting into (2.23) and expanding to linear order in ω

$$\eta_{\mu\nu} = (\delta^{\alpha}_{\mu} + \omega^{\alpha}_{\mu})\eta_{\alpha\beta}(\delta^{\beta}_{\nu} + \omega^{\beta}_{\ \nu}) = \eta_{\mu\nu} + (\omega_{\nu\mu} + \omega_{\mu\nu}) + O(\omega^2),$$

and then

$$\omega_{\mu\nu} = -\omega_{\nu\mu}. \tag{2.28}$$

That is, $\omega_{\mu\nu}$ is antisymmetric on its indices. The space of real antisymmetric 4×4 matrices is 6 dimensional, so the Lorentz group is 6 dimensional.

Now we will parametrize $\omega_{\mu\nu}$ using a basis of 6 antisymmetric matrices 4×4 $M^{\alpha\beta}$ $(\alpha, \beta = 0, \cdots, 3)$

$$\omega_{\mu\nu} = \sum_{\alpha<\beta} \omega_{\alpha\beta} \left\{ (M^{\alpha\beta})_{\mu\nu} \right\} \tag{2.29}$$

where $\omega_{\alpha\beta}$ are real numbers.

The infinitesimal transformation (2.27) can be written as

$$\Lambda^{\mu}_{\ \nu} = \delta^{\mu}_{\nu} + \omega^{\alpha}_{\ \beta}(M^{\alpha\beta})^{\mu}_{\nu}. \tag{2.30}$$

To make contact with the more familiar generators of rotations and boost, it is convenient to define

$$\begin{aligned} J_i &\equiv -\frac{1}{2}\epsilon_{ijk}M^{jk} \\ K_i &\equiv M^{0i} \end{aligned} \tag{2.31}$$

where ϵ_{ijk} is the totally antisymmetric symbol. They satisfy the commutation relations

$$\begin{aligned} [J_i, J_j] &= \epsilon_{ijk}J_k, \\ [J_i, K_j] &= \epsilon_{ijk}K_k, \\ [K_i, K_j] &= -\epsilon^{ijk}J_k. \end{aligned} \tag{2.32}$$

The more general Lorentz transformation has the form

$$A(\mathbf{n}, \mathbf{m}) = \exp\{\ (-\mathbf{n}\cdot\mathbf{J} - \mathbf{m}\cdot\mathbf{K})\}, \tag{2.33}$$

where $\mathbf{n}$ and $\mathbf{m}$ are the 6 parameters of the group. J^i are the rotation generators and K^i are the boost generators.

COMMENTS

In quantum mechanics the commutations relations (2.32) change to

$$\begin{aligned} [J_i, J_j] &= i\hbar\epsilon_{ijk}J_k, \\ [J_i, K_j] &= i\hbar\epsilon_{ijk}K_k, \\ [K_i, K_j] &= -i\hbar\epsilon_{ijk}J_k, \end{aligned} \tag{2.34}$$

and the more general Lorentz transformation has the form

$$U(\mathbf{n},\mathbf{m}) = \exp -\frac{i}{\hbar}(\mathbf{n}\cdot\mathbf{J} + \mathbf{m}\cdot\mathbf{K}). \tag{2.35}$$

2.2.2 Universal covering of the Lorentz group

Let $H(2,\mathbb{C})$ be the space of 2×2 Hermitian matrices $A(A^T = \bar{A})$. A basis for $H(2,\mathbb{C})$ is given by the identity matrix and the three Pauli matrices σ^i. There is an isomorphism $\mathbb{R}^4 \to H(2,\mathbb{C})$ given by

$$x \to \widetilde{x} \equiv x_0\mathbf{1} + \sigma\mathbf{x} \tag{2.36}$$

Recall that $SL(2,\mathbb{C})$ is the group of 2×2 complex matrix A with $\det A = 1$.

Theorem 2.2.2. *There is a homomorphism $F : SL(2,\mathbb{C}) \to L_+^\uparrow$ given by $(F(A)(x)) = A\widetilde{x}A^\dagger$ where $A^\dagger = \bar{A}^\top$, $A \in SL(2,\mathbb{C})$ and $x \in \mathbb{R}^4$. Moreover, F is onto with $F^{-1}(I) = \pm I$. F is the so-called universal covering homomorphism.*

Definition 2.2.3. We define a representation $\rho :\ SL(2,\mathbb{C}) \to GL(4,\mathbb{C})$ by

$$\rho(A) = \begin{bmatrix} A & 0 \\ 0 & A^{\dagger -1} \end{bmatrix}, \quad A \in SL(2,\mathbb{C}). \tag{2.37}$$

Clearly ρ is the direct sum of two irreducible representations, commonly denoted by $D^{(1/2,0)} : SL(2,\mathbb{C}) \to GL(2,\mathbb{C})$ and $D^{(0,1/2)} : SL(2,\mathbb{C}) \to GL(2,\mathbb{C})$, given by $D^{(1/2,0)}(A) = A$ and $D^{(0,1/2)}(A) = A^{\dagger -1}$.

Theorem 2.2.4. *The representations $D^{(1/2),0)}$ and $D^{(0,1/2)}$ are not equivalent.*

2.3 Differential forms

Definition 2.3.1. A *curve* through a point $x \in M$ is a map $\gamma : (a,b) \to M\,(a < 0 < b)$ such that $\gamma(0) = x$. Curves γ_1 and γ_2 through x are called *equivalent* if $(\varphi\circ\gamma_1)'(0) = (\varphi\circ\gamma_2)'(0)$ for some chart $\varphi : U \to \mathbb{R}^n$ with $x \in U$. An equivalent class of curves through x is called a *tangent vector at x*; the set of all tangent vectors at x is denoted by T_xM. We write $\gamma'(0)$ for the vector in T_xM determined by γ. Note that T_xM has a natural vector space. If $Y_x \in T_xM$ (say $Y_x = \gamma'(0)$) and $f \in C^\infty(M)$, then $(f\circ\gamma)'(0) \in \mathbb{R}$ is called *derivative of f along Y_x*, and is denoted by $Y_x[f]$.

Definition 2.3.2. Let $TM = \cup_{x\in M} T_xM$. A *vector field* on M is a function $Y :\ M \to TM$ such that $Y_x \in T_xM$ and for all $f \in C^\infty(M)$ the function $x \to Y_x[f]$ is in $C^\infty(M)$; we denote this function by $Y[f]$. The set of all vector fields on M is denoted by $\Gamma(TM)$. If $Y, Z \in \Gamma(TM)$, then $[Y,Z]$ is that vector field such that $[Y,Z]_x[f] = Y_x[Z[f]] - Z_x[Y[f]]$.

Definition 2.3.3. If $f : M \to N$ is a map and $x \in M$, then $f_{*x} : T_xM \to T_{f(x)}N$ is the linear function (*differential of f at x*) $f_{*x}(\gamma'(0)) = (f\circ\gamma)'(0)$ where γ is a curve through x.

Definition 2.3.4. Let $\varphi : U \to \mathbb{R}^n$ be a chart. The *coordinate vector fields* $\partial_1, \cdots, \partial_n$ on $U \subset M$ are defined by

$$(\partial_i)_x = \frac{d}{dt}\varphi^{-1}(\varphi(x) + te_i)|_{t=0}$$

where e_i is the standard unit vector in $\mathbb{R}^n$. Any $Y \in \Gamma(TM)$, when restricted to U, can be expressed as $Y = \sum a^i\partial_i$ where $a^i \in C^\infty(U)$.

Definition 2.3.5. Let $Y \in \Gamma(TM)$ be such that (for each $x \in M$) there is curve $\gamma_x : \mathbb{R} \to M$ through x with $\gamma_x'(t) = Y_{\gamma_x}(t)$. For $t \in \mathbb{R}$, define $\varphi_t : M \to M$ by $\varphi_t(x) = \gamma_x(t)$. We can prove that φ_t is a diffeomorphism, and that $\varphi_s \circ \varphi_t = \varphi_{s+t}$ for all $s, t \in \mathbb{R}$. The set $\{\varphi_t | \in \mathbb{R}\}$ is called the *one-parameter group generated by* Y. If $Z \in \Gamma(TM)$, then the *Lie derivative of* Z *along* Y is the vector field $L_Y Z$ defined by

$$L_Y Z = \frac{d}{dt}\varphi_{t*}^{-1}(Z)|_{t=0}$$

We can prove that $L_Y Z = [Y, Z]$. For a chart $\varphi : U \to \mathbb{R}^n$ with $Y = \sum a^i \partial_i$ and $Z = \sum b^j \partial_j$ we have $[Y, Z] = \sum (a^i \partial_i [b^j] - \sum b^i \partial_i [a^j] \partial_j = L_Y Z$ on U.

Definition 2.3.6. The space of all tensor fields of type (p, q) on M is denoted by $T^{p,q}(M)$ such that $\omega_x \in \Delta^k(T_x M)$.

Definition 2.3.7. A *k-form* on M is a tensor field $W \in T^{0,k}$. The space of k-forms on M is denoted by $\delta^k(M)$. For $\alpha \in \Delta^i(M)$ and $\beta \in \Delta^j(M)$, we define $\alpha \wedge \beta \in \Delta^{i+j}(M)$ by $(\alpha \wedge \beta)_x = \alpha_x \wedge \beta_x$. If $\phi : U \to \mathbb{R}^n$ is a chart $\phi = (x^1, \ldots, x^n)$ $(x^i \in C^\infty(U))$, then $dx^1, \ldots, dx^n$ are defined to be those 1-forms with $dx^i(\delta_j) = \delta^i_j$. Any $\omega \in \Delta^k(M)$ can be written on U as

$$\omega = \frac{1}{k!} \sum \omega_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}, \tag{2.38}$$

where $\omega_{i_1 \cdots i_k} \in C^\infty(U)$.

Definition 2.3.8. For $\omega \in \Delta^k(M)$, we define $d\omega$ to be the $(k+1)$-form that when restricted to U is given by

$$d\omega = \frac{1}{k!} \sum d(\omega_{i_k \cdots i_k}) dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \frac{1}{k!} \sum \partial_i \left\{\omega_{i_k \cdots i_k}\right\} dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}. \tag{2.39}$$

In fact, $d\omega$ can be defined as that $(k+1)$-form. The operation $d : \Delta^k(M) \to \Delta^{k+1}(M)$ is called *exterior differentiation*. If $\alpha \in \Delta^i(M)$ and $\beta \in \Delta^j(M)$, then $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^i \alpha \wedge d\beta$; and $d^2 \equiv d \circ d = 0$.

Definition 2.3.9. . If $f : M \to N$ is a map and $\omega \in \Delta^k(N)$, the the *pull-back* $f^*\omega \in \Delta^k(M)$ is defined by $(f^*\omega)_x(Y_1, \ldots, Y_k) = \omega_{f(x)}(f_{*x}Y_1, \ldots, f_{*x}Y_k)$ for $Y_1, \ldots Y_k \in T_x M$. When $k = 0$, $f^*\omega \equiv \omega \circ f \in C^\infty(M)$. It can be proved that $df^*\omega = f^* d\omega$, $f^*(\alpha \wedge \beta) = f^*\alpha \wedge f^*\beta$, and $(f \circ g)^*\omega = g^* f^* \omega$.

2.3.1 Maxwell's equations and differencial forms

Let $M = \mathbb{R}^4$ with coordinates $\left(x^0, x^1, x^2, x^3\right) = (t.x, y, z))$ and metric g such that $g\left(\partial_0, \partial_0\right) = 1$, $g\left(\partial_i, \partial_i\right) = -1$ for $i = 1, 2, 3$ and $g\left(\partial_i, \partial_j\right) = 0$ for $i \neq j$ (i.e., (M, g) is a *Minkowski space*. Considerer the 2-form $F = E_1 dx \wedge dt + E_2 dy \wedge dt + E_3 dz \wedge dt + B_1 dy \wedge dz + B_2 dz \wedge dy + B_3 dx \wedge dy$. For $d\mathbf{r} = (dx, dy, dz)$ and $d\sigma = (dy \wedge dz, dz \wedge dx, dx \wedge dy)$, we employ the shortland $F = \mathbf{E} \cdot d\mathbf{r} \wedge dt + \mathbf{B} \cdot d\sigma$. By a simple computation, we obtain

$$dF = \left(\nabla \times \mathbf{E} + \frac{\partial B}{\partial t}\right) d\sigma \wedge dt + (\nabla \cdot \mathbf{B})\, d\tau$$

where $d\tau = dx \wedge dy \wedge dz$. Thus, $dF = 0$ iff $\nabla \times \mathbf{E} + \dfrac{\partial B}{\partial t} = 0$ and $\nabla \cdot \mathbf{B} = 0$, which are two of Maxwell's equations (where $\mathbf{E}$ is the *electric field* and $\mathbf{B}$ is the *magnetic field*). Now,

$*F = \mathbf{E}\cdot d\sigma - \mathbf{B}\cdot d\mathbf{r} \wedge dt$, and so

$$d * F = (\nabla\cdot\mathbf{E})\, d\tau - \left(\nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t}\right)\cdot d\sigma \wedge dt.$$

Now $\delta = -(-1)^g(-1)^{4(k+1)} * d * = * d *$ on $\Lambda^k\left(\mathbb{R}^4\right)$. Thus, $\delta = * d * F = (\nabla\cdot\mathbf{E})\, dt - \left(\nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t}\right)\cdot d\mathbf{r}$. Let the maps $\rho : \mathbb{R}^4 \to \mathbb{R}$ and $\mathbf{J} : \mathbb{R}^4 \to \mathbb{R}^3$ be the *charge density* and the current density, rexpectively. Then $\delta = j$ is equivalent to the other two (inhomogeneous) Maxwell equations, $\nabla\cdot\mathbf{E} = \rho$ and $\nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t} = \mathbf{J}$. Thus, the four Maxwell equations are summarized by

$$\begin{aligned} dF &= 0 \\ \delta F &= j. \end{aligned} \tag{2.40}$$

Applying δ to $\delta F = j$, we obtain

$$\begin{aligned} 0 = \delta^2 F = \delta j = * d * j = * d \left(\rho d\tau - \mathbf{J}\cdot d\sigma \wedge dt\right) \\ = * \left(\frac{\partial\rho}{\partial t} dt \wedge d\tau - \nabla\cdot\mathbf{J} d\tau \wedge dt\right) = -\left(\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J}\right). \end{aligned}$$

Thus, we obtain the *continuity equation*

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{J} = 0 \tag{2.41}$$

2.4 Principal fiber bundles, connections and curvature

2.4.1 Principal bundles and connections

Definition 2.4.1. A *principal fiber bundle* (PFB) consists of a manifold P, a Lie group G, a base manifold M, and a projection map $\pi : P \to M$ such that (A), (B), and (C) following hold.

(A) For each $g \in G$ there is a diffeomorphism $R_g : P \to P$ (we write $R_g(p) = pg$) such that $p(g_1 g_2) = (pg_1)g_2$ for all $g_1, g_2 \in G$ and $p \in P$; and if $e \in G$ is the identity element, then $pe = p$ for all $p \in P$. We require the function $P \times G \to P$ given by $(p, g) \mapsto pg$ to be a map. We suppose that if $pg = p$ for some $p \in P$ and $g \in G$, then $g = e$.

(B) The map $\pi : P \to M$ is onto, and $\pi^{-1}(\pi(p)) = \{pg : g \in G\}$ (which is, by definition, the orbit of G through p). If $x \in M$, then $\pi^{-1}(x)$ is called the *fiber* above x.

(C) For each $x \in M$ there is an open set U with $x \in U$ and a diffeomorphism $T_u : \pi^{-1}(U) \to U \times G$ of the form $T_u(p) = (\pi(p), s_u(p))$ where $s_u : \pi^{-1}(U) \to G$ has the property $s_u(pg) = s_u(p)g$ for all $g \in G$, $p \in \pi^{-1}(U)$. The map T_u is called a *local trivilization* (LT), or (in physics language) a *choice of gauge*.

EXAMPLE

Let M be an n-manifold. We will define a PFB $\pi : L(M) \to M$ with group $GL(n, \mathbb{R})$, called the *frame bundle of M*. A *frame* at $x \in M$ is a linear isomorphism $u : \mathbb{R}^n \to T_x M$. Note that such a frame determines a basis $u(e_1), \cdots, u(e_n)$ of $T_x M$ where $e_1, \cdots, e_n$ is the usual basis of $\mathbb{R}^n$. Let $L(M)_x$ be the set of all frames at x, and set

$$L(M) = \bigcup_{x \in M} L(M)_x.$$

For $u \in L(M)_x$, $\pi(u) \equiv x$. For $A \in GL(n,\mathbb{R})$, we define $R_A : L(M) \to L(M)$ by $R_A(u) = u \circ A$. This a free, right action of $GL(n,\mathbb{R})$ on $L(M)$, but we need to put a differentiable structure on $L(M)$ before we can speak of $\pi : L(M) \to M$ as being C^∞. Let $W \subset M$ be a coordinate neighborhood with coordinates $x^1, \cdots, x^n$ and associated coordinates fields $\partial_1, \cdots, \partial_n$ on W. Define a map $\sigma : W \to L(M)$ by lettitng $\sigma(y) : \mathbb{R}^n \to T_yM$ be the isomorphism such that $\sigma(y) = (\partial_i)_y$.

Definition 2.4.2. Let $\mathfrak{g}$ be the Lie algebra of G. A *connection* is a $\mathfrak{g}$-valued 1-form ω defined on P such that the properties (a) and (b) hold.

(a) Let $A \in \mathfrak{G}$ and let A^Λ be the vector field on P defined by

$$A^\Lambda_p = \frac{d}{dt}\left(p \exp(tA)\right)|_{t=0}.$$

Then $\omega(A^\Lambda_p) = A$. A^Λ is called *fundamental field.*

(b) For $g \in G$, let . We call ω a *connection 1-form.*

Physicist refer to the 1-form ω_u as *gauge potentials.*

2.4.2 Curvature

Given a connection 1-form ω on a PFB $\pi : P \to M$ with group G, we can write any $X \in T_pP$ as $X = X^V + X^H$ where X^V is vertical (i.e., $\pi_*(X^V) = 0$) and X^H is horizontal (i.e. $\omega(X^H) = 0$).

Definition 2.4.3. If $\varphi \in \Lambda^k(P,\mathfrak{g})$, then we define $\varphi^H \in \Lambda^k(P,\mathfrak{g})$ by $\varphi^H(X_1, \cdots, X_k) = \varphi(X_1^H, \cdots, X_k^H)$.

Definition 2.4.4. The *exterior covariant derivative* of $\varphi \in \Lambda^k(P,\mathfrak{g})$ is $D^\omega \equiv (/d\varphi)^H \in \Lambda^{k+1}(P,\mathfrak{g})$ where $d\varphi$ is the usual exterior derivative of φ. Although the operator D^ω depends on ω, it is customary to omit the subscript ω.

Definition 2.4.5. The *curvature of the connection* $\omega \in \Delta^1(P,\mathfrak{g})$ is $\Omega^\omega \equiv D^\omega\omega \in \Delta^2(P,\mathfrak{g})$. When ω is regarded as a potential, Ω^ω is called the *field strength* of ω.

Theorem 2.4.6. *The curvature form is given by* $\Omega^\omega = d\omega + \frac{1}{2}[\omega,\omega]$.

Theorem 2.4.7. *If* ω *is a connection 1-form on* P *with curvature* Ω^ω*, then* $D^\omega\Omega^\omega = 0$ *(Bianchi identity or homogeneous field equation). In fact, we have* $d\Omega^\omega = [\omega^\omega, \omega]$.

Theorem 2.4.8. *If* G *is a matrix group, then* $\Omega^\omega = dw + w \wedge w$ *and* $\Omega_u = d\omega_u + \omega_u \wedge \omega_\mu$.

2.5 Lagrangians, particle fields and gauge transformations

A particle field can be regarded as a section of a vector bundle associated to some PFB, or equivalently, as a vector-valued function on P with certain transformation properties. The particle field obeys a differential equation (Euler-Lagrange's equation) obtained by setting the first variation of the integral of the Lagrangian density (principle of least Action) equal o zero.

Gauge transformations are defined as base-preserving automorphisms of the PFB. Locally, they amount to a variable change of *internal* reference frame or gauge. There is a little hope that

the Lagrangian is physically meaningful unless it is invariant under gauge transformations of the particle field. However, we will find that this invariance is not possible without introducing gauge potentials into the Lagrangian by replacing the ordinary differential of the particle field by its covariant derivative. The gauge potentials (particularly after quantization) can be interpreted as various forms of radiation (photons, gluons, intermediate vector bosons, etc.).

2.5.1 Particle fields

Let $\pi: P \to M$ be a PFB with group G. Suppose that G acts on some manifold M to the left. That is, for each $g \in G$, there is a map $L_g : M \to M$ (we write $g \cdot f = L_g(f)$) such that $e \cdot f = f$, $(g_1 g_2) \cdot f = g_1 \cdot (g_2 \cdot f)$, and the map $G \times M \to M$, given by $(g, f) \to g \cdot f$, is C^∞. If F is a vector space V and $L_g : V \to V$ is linear, then the homomorphism $G \to GL(V)$ given by $g \to L_g$ is called *representation* of G. Two representations are $G \to GL(V)$ and $G \to GL(V')$, say $g \to L_g$ and $g \to L'_g$, respectively, are *equivalent* if there is a linear isomorphism $T: V \to V'$ such that $L'_g = T \circ L_g \circ T^{-1}$ for all g is G. We define $C(P, M)$ to be the space of all maps $\tau(pg) \to g^{-1} \cdot \tau(p)$. In the case where the action of G defines a representation $G \to GL(V)$, the elements of $C(P, V)$ are called *particle fields.*

2.5.2 Gauge transformations

An automorphism of a PFB $\pi : P \to M$ is a diffeomorphism $f : P \to P$ such that $f(pg) = f(p)g$ for all $g \in G$, $p \in P$. Note that f induces a well-defined diffeomorphism $\bar{f} : M \to M$ given by $\bar{f}(\pi(p)) = \pi(f(p))$. A *gauge transformation* of a PFB is an automorphism $f : P \to P$ such that $\bar{f} = 1_M$ (i.e., $\pi(p) = \pi(f(p))$). We set $GA(P)$ the group of gauge transformations.

If $\pi : M \times G \to M$ is a product bundle, then the maps $L_g : M \times G \to M \times G$ given by $L_g(x, g') = (x, gg')$ are in $GA\,(M \times G)$. Physicist call such transformations *global* since g does not depend on x. *Local* gauge transformations are those of the form $(x, g') \to (x, h(x)g')$ where $h : M \to G$ is not necessarily constant.

2.5.3 Lagrangians and Euler-Lagrange equations

Definition 2.5.1. A *Lagrangian* on a manifold M is a smooth function $L : TM \to \mathbb{R}$. For any smooth curve $\gamma : [a, b] \to M$, define

$$I_L(\gamma) = \int_a^b L(\dot{\gamma}(t))dt. \tag{2.42}$$

I_L is called the *functional* associated to L.

Given a curve $\gamma : [a, b] \to M$, a (smooth) variation of γ with fixed endpoints is, by definition, a smooth map

$$\Gamma : [a, b] \times (-\varepsilon, \varepsilon) \to M \tag{2.43}$$

for some $\varepsilon > 0$ with the property that $\Gamma(t, 0) = \gamma(t)$ for all $t \in [a, b]$ and that $\Gamma(a, s) = \gamma(a)$ and $\Gamma(b, s) = \gamma(b)$ for all $s \in (-\varepsilon, \varepsilon)$.

If $\mathcal{L}$ is a Lagrangian on M and Γ is a variation of $\gamma : [a, b] \to M$, then we can define a function $I_{L,\Gamma} : (-\varepsilon, \varepsilon) \to \mathbb{R}$ by setting

$$I_{L,\Gamma}(s) = I_L(\gamma_s) \tag{2.44}$$

where $\gamma_s(t) = \Gamma(t, s)$.

Definition 2.5.2. A curve $\gamma : [a, b] \to M$ is *L-critical* if $I'_{L,\Gamma}(0) = 0$ for all variations of γ.

Definition 2.5.3. If $U \subset M$ is an open set on which there is a coordinate chart $x : U \to \mathbb{R}^n$, then there is a canonical extension of these coordinates to a coordinate chart $(x, p) : TU \to \mathbb{R}^n \times \mathbb{R}^n$ with the property that, for any curve $\gamma : [a, b] \to U$, with coordinates $y = x \circ \gamma$, the p-coordinates of the curve $\dot{\gamma} : [a, b] \to TU$ are given by $p \circ \dot{\gamma} = \dot{y}$. We shall call the coordinates (x, p) on TU, the *canonical coordinates* associated to the coordinate system x on U.

In a canonical coordiante system (x, p) on TU where U is an open set in M, the function L can be expressed as a function $\mathcal{L}(x, p)$ of x and p. For a curve $\gamma : [a, b] \to M$ which happens to lie in U, the functional I_L becomes simply

$$I_L = \int_a^b L(y(t), \dot{y}(t))\, dt. \tag{2.45}$$

We will now derive the classical conditions for such a γ to be L-critical: Let $h : [a, b] \to \mathbb{R}^n$ be any smooth map which satisfies $h(a) = h(b) = 0$. Then, for sufficiently small ε, there is a variation Γ of γ which is expressed in (x, p)-coordinates as

$$(x, p) \circ \Gamma = (y + sh, \dot{y} + s\dot{h}). \tag{2.46}$$

The, by the classic integration-by-parts method,

$$\begin{aligned} I'_{L,\Gamma}(0) &= \frac{d}{ds}\bigg|_{s=0} \left(\int_a^b L(y(t) + sh(t), \dot{y}(t) + s\dot{h}(t))\, dt \right) \\ &= \int_a^b \left[\frac{\partial L}{\partial x^k}(y(t), \dot{y}(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial p^k}(y(t), \dot{y}(t)) \right) \right] h^k(t)\, dt. \end{aligned} \tag{2.47}$$

This formula is valid for any $h : [a, b] \to \mathbb{R}^n$ which vanishes at the endpoints. It follows without difficulty that the curve γ is L-critical if only if $y = x \circ \gamma$ satisfies the n differenctial equations

$$\frac{\partial L}{\partial x^k}(y(t), \dot{y}(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial p^k}(y(t), \dot{y}(t)) \right) = 0 \quad \text{for } 1 \le k \le n. \tag{2.48}$$

These are the famous *Euler-Lagrange equations.*

Now, we generalize the Euler-Lagrange equations to include classical field theories with an infinite number of degrees of freedom. We begin with a Lagrangian density that is a function of both the field $\phi(x)$ as well as its space time derivatives $\partial_\mu \phi(x)$:

$$L(\phi(x), \partial_\mu \phi(x)) \tag{2.49}$$

The action is given by a four dimensional integral over a Lagrangian density $\mathcal{L}$:

$$\begin{aligned} L &= \int d^3x\, \mathcal{L}(\psi, \partial_\mu \phi) \\ I &= \int d^4x\, \mathcal{L} = \int dt\, L \end{aligned} \tag{2.50}$$

integrated between initial and final times.

As before, we can retrieve the classical equations of motion by minimizing the action and we obtain:

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} - \frac{\delta \mathcal{L}}{\delta \phi} = 0 \tag{2.51}$$

2.5.4 Noether's theorem

One of the achievements of the lagrangian formalism is that we can use the symmetries of the action to derive conservation principles. For example, in classical mechanics, if the Hamiltonian is time independent, then energy is conserved. Likewise, if the Hamiltonian is translation invariant in three dimensions, then momentum is conserved. And if the Hamiltonian is rotational invariant in three dimensions, then angular momentum is conserved.

The precise mathematical formulation of this correspondence is given by the Noether's theorem. In general, an action may be invariant under either internal, isospin symmetry transformation of the fields, or under some space-time symmetry. We will only discuss the isospin symmetry, where the fields ϕ^α vary according to some small parameter $\delta\epsilon^\alpha$. The action varies as

$$\delta I = \int d^4x \left(\frac{\delta \mathcal{L}}{\delta \phi^\alpha} \delta\phi^\alpha + \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi^\alpha} \delta \partial_\mu \phi^\alpha \right) = \int d^4x \, \partial_\mu \left(\frac{\delta \mathcal{L}}{\delta \partial_\mu \phi^\alpha} \delta\phi^\alpha \right) \tag{2.52}$$

where we have used the equations of motion and have converted the variation the action into the integral of a total derivative. This defines the *current* J^μ_α,

$$J^\mu_\alpha = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi^\beta} \frac{\delta \phi^\beta}{\delta \epsilon^\alpha} \tag{2.53}$$

If the action is invariant under this transformation, then we have established that the current is conserved,

$$\partial_\mu J^\mu_\alpha = 0 \tag{2.54}$$

From this current, we can also establish a conserved charge, given by the integral over the fourth component of the current

$$Q_\alpha \equiv \int d^3x \, J^0_\alpha \tag{2.55}$$

Now let us integrate the conservation equation:

$$\begin{aligned} 0 &= \int d^3x \, \partial_\mu J^\mu_\alpha = \int d^3x \, \partial_0 J^0_\alpha + \int d^3x \, \partial_i J^i_\alpha \\ &= \frac{d}{dt} \int d^3x \, J^0_\alpha + \int dS_i \, J^i_\alpha = \frac{d}{dt} Q_\alpha + \text{surface term} \end{aligned} \tag{2.56}$$

Let us assume that the fields appearing in the surface term vanish sufficiently rapidly at infinity so that the last term can be neglected. Then:

$$\partial_\mu J^\mu_\alpha = 0 \rightarrow \frac{d\, Q_\alpha(t)}{dt} = 0 \tag{2.57}$$

In summary, the symmetry of the action implies the conservation of a current J^μ_α, which in turn implies a conservation principle:

Symmetry $\rightarrow$ Current conservation $\rightarrow$ Conservation principle

Chapter 3

The beginnings of gauge theory

3.1 Gauge theory in classical electrodynamics

Maxwell's equations (in Heaviside-Lorentz rationalized units) are

$$\begin{aligned} &\text{(a) } \nabla\cdot\mathbf{B} = 0, \qquad &&\text{(b) } \nabla\times\mathbf{E} + \frac{\partial\mathbf{B}}{\partial t} = 0, \\ &\text{(c) } \nabla\cdot\mathbf{E} = \rho, \qquad &&\text{(d) } \nabla\times\mathbf{B} - \frac{\partial\mathbf{E}}{\partial t} = \mathbf{j}, \end{aligned} \tag{3.1}$$

(a) tells us there are no magnetic monopoles. (b) is Faraday's law; a changing magnetic field produces an electric field. (c) is Gauss's law; the total charge inside a closed surface may be obtained by integrating the normal component of $\mathbf{E}$ over the surface. (d) is Ampere's generalized law, stating that changing electric fields or currents produce magnetic fields. The equations (a) and (b) are known as the homogeneous equations, (c) and (d) as the inhomogeneous ones.

Introducing the 4-vector potencial

$$A^\mu = (\phi, \mathbf{A}) \tag{3.2}$$

with

$$\mathbf{B} = \nabla\times\mathbf{A}, \quad \mathbf{E} = -\frac{\partial\mathbf{A}}{\partial t} - \nabla\phi, \tag{3.3}$$

equations (a) and (b) are automatically satisfied. The electric and magnetic fields are the components of a tensorial field defined by

$$F^{\mu\nu} = -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{3.4}$$

where (recall that $\partial^i = -\partial_i$)

$$\begin{aligned} F^{0i} &= \partial^0 A^i - \partial^i A^0 = -E^i \\ F^{ij} &= \partial^i A^j - \partial^j A^i = -\varepsilon^{ijk} B^k, \end{aligned} \tag{3.5}$$

where $\varepsilon^{ijk} = \varepsilon_{ijk}$ is the totally antisymmetric Levi-Civita symbol. Equations (3.5) may be displayed in matrix form, with the rows and columns corresponding to the numbers 0,1,2,3:

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E^1 & -E^2 & -E^3 \\ E^1 & 0 & -B^3 & B^2 \\ E^2 & B^3 & 0 & -B^1 \\ E^3 & -B^2 & B^1 & 0 \end{pmatrix} \tag{3.6}$$

$F^{\mu\nu}$ is called the *electromagnetic field tensor*. It transforms, under Lorentz transformations, like an antisymmetric second rank tensor:

$$F^{\alpha\beta} = \Lambda^{\alpha}_{\ \mu}\Lambda^{\beta}_{\ \nu}F^{\mu\nu}$$

Now consider the inhomogeneous equations. It is straightforward to verify that they are both contained in the covariant equation

$$\partial_\mu F^{\mu\nu} = j^\nu \tag{3.7}$$

with

$$j^\nu = (\rho, \mathbf{j}) \tag{3.8}$$

It is easy to verify that

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0 \tag{3.9}$$

Now we define the *dual tensor* $\widetilde{F}^{\mu\nu}$ by

$$\widetilde{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma} \tag{3.10}$$

where $\varepsilon^{\mu\nu\rho\sigma}$ is the Levi-Civita symbol in four dimensions (with $\varepsilon^{0123} = 1$). Its elements are

$$\widetilde{F}^{\mu\nu} = \begin{pmatrix} 0 & -B^1 & -B^2 & -B^3 \\ B^1 & 0 & E^3 & -E^2 \\ B^2 & -E^3 & 0 & E^1 \\ B^3 & E^2 & -E^1 & 0 \end{pmatrix} \tag{3.11}$$

Because of the antisymmetry of $\varepsilon^{\mu\nu\rho\sigma}$ and (3.9), it follows the equation

$$\partial_\mu \widetilde{F}^{\mu\nu} = 0 \tag{3.12}$$

Equations (3.7) and (3.12) are the Maxwell's equations in covariant form.

Under a transformation (*gauge transformation*)

$$A^\mu \to A^\mu + \partial^\mu \chi, \tag{3.13}$$

where χ is an arbitrary scalar function, $\mathbf{E}$ and $\mathbf{B}$ (and $F^{\mu\nu}$) remain unchanged. We can see that A^μ satisfies

$$\Box A^\nu - \partial^\nu \left(\partial_\mu A^\mu\right) = j^\nu \tag{3.14}$$

We may make use of the freedom (3.13) and choose a particular χ so that the transformed A^μ satisfies the *Lorentz gauge condition*

$$\partial_\mu A^\mu = 0 \tag{3.15}$$

In this gauge and without sources ($j^\mu = 0$) equation (3.14) becomes

$$\Box A^\mu = 0, \tag{3.15}$$

which means that the electromagnetic field, when its quantum nature is considered, will be seen to correspond to massless particles which travel at the speed of light.

If we generalize Maxwell's equations to massive particles (m is the mass) we have the *Proca equations*:

$$\begin{gathered} F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \\ \partial_\mu F^{\mu\nu} + m^2 A^\nu = 0. \end{gathered} \tag{3.16}$$

Taking the divergence we have

$$m^2 \partial_\nu A^\nu = 0, \tag{3.17}$$

and since $m^2 \neq 0$, we find $\partial_\nu A^\nu = 0$, and we have lost the freedom of gauge transformations which Maxwell's equations had. We conclude that the Proca equations are not gauge invariant.

3.2 Gauge theory in general relativity

3.2.1 Einstein's field equation

Einstein's original purpose was to explain the equivalence of gravitational and inertial mass, but in doing so he revolutionized the theory by showing that it could be attributed entirely to the geometry of space-time. The Einstein's gravitational theory, which was based on Riemannian geometry, also provided the inspiration for the non-Riemannian geometry which is at the heart of gauge theory.

The mathematician Levi-Civita introduced the concept of *parallel transfer*. What Levi-Civita realized was, in fact, the covariance of the Riemannian derivative and the Riemann tensor. A connection is defined as an array of functions $\Gamma^{\lambda}_{\mu\nu}(x)$ (at first symmetric in the lower two index) which satisfies the transformation law

$$\Gamma'^{\gamma}_{\alpha\beta}(x') = \frac{\partial x^{\mu}}{\partial x'^{\alpha}} \frac{\partial x^{\nu}}{\partial x'^{\beta}} \frac{\partial x'^{\gamma}}{\partial x^{\lambda}} \Gamma^{\lambda}_{\mu\nu}(x) + \frac{\partial^2 x^{\mu}}{\partial x'^{\alpha} \partial x'^{\beta}} \frac{\partial x'^{\gamma}}{\partial x^{\mu}} \tag{3.18}$$

Each connection Γ defines a *covariant derivative* ∇_{μ}

$$(\nabla_{\mu})^{\alpha}_{\beta} = \delta^{\alpha}_{\beta}\partial_{\mu} + \Gamma^{\alpha}_{\mu\beta} \tag{3.19}$$

and each such ∇_{μ} defines a Riemann tensor

$$R^{\alpha}_{\mu\nu\beta} = [\nabla_{\mu}, \nabla_{\nu}]^{\alpha}_{\beta} \tag{3.20}$$

The Riemann tensor is antysymmetric with respect to the indices μ and ν and the sum of the tensors obtained by a cyclic permutation of the lower indices is zero (Bianchi identity). These are the symmetries in absence of metric. When metric is brought onto the scene it impresses on Riemann the additional symmetry

$$R_{\alpha\beta\gamma\delta} = -R_{\beta\alpha\gamma\delta} \quad \text{(antisymmetry on first two indices)}. \tag{3.21}$$

The symmetries reduce (in four dimensions) the number of independent components from 256 to 20.

Behind this algebraic formalism lay the geometrical idea of parallel transfer, by which a vector $v(x)$ is transferred along the curve C using not the infinitessimal increments $dv = (\partial_{\mu} v)\, dx^{\mu}$, but infinitesimal increments of the form

$$\delta v = (\nabla_{\mu} v)\, dx^{\mu}, \tag{3.22}$$

where the dx^{μ} are tangent to the curve and ∇_{μ} is the covariant derivative defined in (3.19).

The transfer (3.22) is independent of the coordinate system because ∇v transform as tensors under general coordinate transformations

$$(\nabla v)^{\alpha}_{\beta}(x') = \frac{\partial x^{\nu}}{\partial x'^{\beta}} \frac{\partial x'^{\alpha}}{\partial x^{\mu}} (\nabla v)^{\mu}_{\nu}(x) \tag{3.23}$$

The *geodesic* may be defined as the curve obtained by the parallel transfer of a vector in its own direction, and (for a suitable choice of the curve parameter τ) the equation of the geodesics then turns out to be

$$\frac{d^2 x^{\lambda}}{d\tau^2} + \Gamma^{\lambda}_{\mu\nu} \frac{dx^{\mu}}{d\tau} \frac{dx^{\nu}}{d\tau} = 0. \tag{3.24}$$

In 1915 Albert Einstein derived "his "field equation

$$G_{\mu\nu} = 8\pi T^{\mu\nu} \tag{3.25}$$

where $T^{\mu\nu}$ is the stress-energy tensor and $G^{\mu\nu}$ is the Einstein curvature tensor,

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R \tag{3.26}$$

with

$$R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu} \qquad R = R^{\mu}{}_{\mu} \tag{3.27}$$

where $R_{\mu\nu}$ is the *Ricci curvature tensor* and R is the *curvature scalar*.

The Bianchi identity has the obvious consequence

$$\nabla_{\beta} G_{\alpha}{}^{\beta} = 0. \tag{3.28}$$

The Einstein's field equation can be derived from Hilbert variational principle, extremalizing the action

$$I = \int \mathcal{L}\, d^4x \tag{3.29}$$

where $\mathcal{L}$ is the Lagrangian density which are divided in two parts,

$$\mathcal{L} = \mathcal{L}_{\text{geom}} + \mathcal{L}_{\text{field}} = \frac{1}{16\pi} g^{\alpha\beta} R_{\alpha\beta} (-g)^{1/2} + \mathcal{L}_{\text{field}} \tag{3.30}$$

In order that the integral I should be an extremun,

$$\delta I = \frac{1}{16\pi} \int \delta \left[g^{\alpha\beta} R_{\alpha\beta} (-g)^{1/2} \right] d^4x + \int \delta \mathcal{L}_{\text{field}} d^4x = 0. \tag{3.31}$$

Considering the Lagrange function L_{field} (scalar function) defined by

$$L_{\text{field}} = \frac{\mathcal{L}_{\text{field}}}{(-g)^{1/2}},$$

we find the Einstein's field equation

$$R_{\alpha\beta} - \frac{1}{2} g_{\alpha\beta} R = 8\pi \left(g_{\alpha\beta} L_{\text{field}} - 2 \frac{\delta L_{\text{field}}}{\delta g^{\alpha\beta}} \right) = 8\pi T_{\alpha\beta}. \tag{3.32}$$

In 1917 Einstein modified field equation introducing the famous *cosmological constant* Λ

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu} \tag{3.33}$$

The modified field equation, by contrast with the original, admits a static, unchanging universe as one particular solution. But when Edwin Hubble discovered the expansion of the universe Einstein abandoned it ("the biggest blunder of my life"). Many workers in cosmology are unwilling to abandon the cosmological constant. Some interpret Λ as stress-energy tensor associated with the vacuum polarization of quantum fields.

3.2.2 Weyl's symmetric generalization

Weyl's point was that standard Riemannian geometry is slightly defective because, although it purports to be infinitesimal, it contains a residue of rigid Euclidean geometry in the fact that the magnitudes of vectors, in contrast to their directions, are path-independent with respect to parallel transport. In Weyl's opinion, a true infinitesimal geometry would not permit this anomaly, and he therefore proposed replacing the Christoffel connection by a (symmetric) connection of the form

$$\Gamma'^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} + \frac{1}{2} g^{\lambda\sigma} \left(g_{\mu\sigma} v_\nu + g_{\sigma\nu} v_\mu - g_{\mu\nu} v_\sigma \right) \tag{3.34}$$

where $v_\mu(x)$ is a vector field. The connection (3.34) is not compatible with the Riemannian metric, and to see what happens, one may take the Riemannian metric at a fixed point x_0 and parallel-transfer it to all (connected) points of the space. One obtains in this way the non-local, non-Riemannian, metric $\tilde{g}_{\mu\nu}$ where

$$\tilde{g}_{\mu\nu} = e^{\int_{x_0}^{x} v_\lambda(y) dy^\lambda} g_{\mu\nu}. \tag{3.35}$$

Since then metric $\tilde{g}$ is used to form inner products, it is easy to see that parallel transfer of a vector from x_1 to x_2 using the connection (3.34) changes the magnitude by a scale factor of the form

$$e^{\int_{x_1}^{x_2} v_\lambda(y) dy^\lambda}. \tag{3.36}$$

The important point is that in general the scale factor is *non-integrable*. That is, it depends on the path taken from x_1 to x_2, and this puts lengths on the same footing as directions. Indeed, the only case in which the scale factor is integrable is when v_μ is a gradient. To see this, one notes that in the path-independent case the integral around a closed loop is zero, and hence by the Stokes theorem,

$$e^{\int f_{\mu\nu} dx^\mu dx^\nu} = e^{\oint v_\mu dx^\mu} = 1 \quad \text{where} \quad f_{\mu\nu} = \partial_\mu v_\nu - \partial_\nu v_\mu, \tag{3.37}$$

where the loop is assumed to be local and therefore not affected by the global topology and the surface-integral is over any 2-dimensional surface tha tspans the loop. Since the loop is arbitrary, this implies that

$$f_{\mu\nu} = 0 \quad \text{or} \quad v_\mu = \partial_\mu \sigma(x) \quad \text{locally}, \tag{3.38}$$

where $\sigma(x)$ is a scalar, as required.

It was in this paper (1918) of Weyl that the word *gauge* was introduced into differential geometry. It was quite appropiate since the scale factor attached to the metric changed the measurement of lenght and the word *gauge* was in common use for measurements of length.

The Weyl geometry was an obvious candidate to describe both gravitation and electromagnetism. Weyl proposed that the geometrical vector v should be identified as a multiple of the electromagnetic field potencial,

$$v_\mu(x) = \frac{e}{\gamma} A_\mu(x), \tag{3.39}$$

where γ is a constant.

Weyl's reasoning was obviously very original and deep but the direct application to gravitational theory turned out to be unnacceptable. The problem is that the lengths of measuring rods and the time measurements of clocks would be rescaled by the non-integrable factor $\exp\left\{\frac{e}{\gamma}\int A_\mu dx^\mu\right\}$ and would therefore depend on their history. This stands in clear contradiction with the fact

that atomic spectra depend only on the nature of the atoms and not on their histories. But the Weyl's general idea was correct in the context of quantum theory. In Weyl's words (1955),

"*I have no doubt that the correct context for the principle of gauge-invariance is here and not, as I believed in 1918, in the intertwining of electromagnetism and gravity.*"

3.2.3 Cartan's theory of gravitation

In 1922 Élie Cartan conjectured that general relativity should be extended by including *torsion*, which allows the Ricci tensor to be non-symmetric. In Cartan's theory the aim was to incorporate not electromagnetism, but (integer) spin in terms of torsion. Although spin-orbit coupling is a relatively minor phenomenom in gravitational physics, Cartan's theory is quite important because it explains the meaning of torsion, which appears naturally in some theories of quantum gravity and it interprets spin as affine torsion, which geometrically is a continuum approximation to a field of dislocations in the space medium.

Cartan considered a $(0,2)$ symmetric tensor $g_{\mu\nu} = g_{\nu\mu}$ called *metric tensor* and an antisymmetric tensor $(1,2)$ tensor $t^{\lambda}_{\mu\nu} = -t^{\lambda}_{\nu\mu}$ known as the *torsion tensor*. Then, Cartan proposed a covariant derivative with two extra properties:

$$\begin{aligned} &a)\ \text{Compatibility of the metric: } \nabla_\lambda g_{\mu\nu} = 0 \\ &b)\ \text{Commutator of the covatiant derivative: } [\nabla_\mu, \nabla_\nu] f = t^{\lambda}_{\nu\mu} \nabla_\lambda f \end{aligned} \tag{3.40}$$

The torsion tensor may also be interpreted as an antisymmetric part of the Christoffel symbols:

$$t^{\lambda}_{\mu\nu} = \Gamma^{\lambda}_{\mu\nu} - \Gamma^{\lambda}_{\nu\mu} \tag{3.41}$$

but such a relation is not manifestly covariant.

The Ricci tensor is no longer symmetric but instead has the property

$$R_{\mu\nu} = R_{\nu\mu} + 3\nabla_{[\mu} t^{\lambda}_{\lambda\nu]} - t^{\lambda}_{\lambda\rho} t^{\rho}_{\mu\nu}. \tag{3.42}$$

Cartan's field equations can be derived by the same variational principle than in general relativity without torsion. Let $\mathcal{L}_M$ represent the Lagrangian density of matter and $\mathcal{L}_G = (-g)^{1/2} R$ represent the Lagrangian density of gravitation. Then, we consider the action

$$I = \int \left((-g)^{1/2} R + 2k\mathcal{L}_M \right) d^4x \tag{3.43}$$

where k $(= 8\pi)$ is the gravitational constant. Variation with respect the metric yiels

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = k P_{\mu\nu} \tag{3.44}$$

where $P_{\mu\nu}$ is the *canonical* energy-momentum tensor. Despite similarity of the form, this is not the same as the Einstein field equation because the Ricci tensor is no longer symemtric but instead contains information about the non-zero torsion tensor as well. The right hand side of the equation cannot be symmetric either, so $P_{\mu\nu}$ must also contain information about the non-zero spin tensor.

We may also consider the variation of I with respect to the torsion tensor $t^{\lambda}_{\mu\nu}$. This yields a new (algebraic) equations

$$t^{\lambda}_{\mu\nu} + g^{\lambda}_{\mu} t^{\rho}_{\nu\rho} - g^{\lambda}_{\nu} t^{\rho}_{\mu\rho} = k s^{\lambda}_{\mu\nu} \tag{3.45}$$

where $s^{\lambda}_{\mu\nu}$ is the *spin tensor*.

The field equations obtained by varying this action with respect to the metric and torsion are

$$G_{\mu\nu} = kT_{\mu\nu} - D_\alpha U^{\alpha}_{\mu\nu} \quad \text{and} \quad T^{\lambda}_{\mu\nu} = kU^{\lambda}_{[\mu\nu]} \tag{3.46}$$

It is well known the problem of the violation, in presence of torsion, of the local gauge invariance of theories gauge due to the straightforward application of the minimal coupling procedure to introduce the interaction with the gauge fields.

In special relativity (we consider Maxwell theory), for the field strength in terms of the gauge potential one has $F = dA = d(A_\nu dx^\nu) = \frac{1}{2}(\partial_\mu A_\nu - \partial_\nu A_\mu)dx^\mu \wedge dx^\nu = F_{\mu\nu}dx^\mu \wedge dx^\nu$. Replacing ∂ by D_μ

$$F'_{\mu\nu} = D_\mu A_\nu - D_\nu A_\mu = \left(\partial_\mu A_\nu - \Gamma^{\lambda}_{\mu\nu}A_\lambda\right) - \left(\partial_\nu A_\mu - \Gamma^{\lambda}_{\nu\mu}A_\lambda\right) = F_{\mu\nu} - 2t^{\lambda}_{\mu\nu}A_\lambda. \tag{3.47}$$

when torsion vanishes, $F'_{\mu\nu} = F_{\mu\nu}$, but when $t^{\lambda}_{\mu\nu} \neq 0$, $F'_{\mu\nu} \neq F_{\mu\nu}$ and the strength tensor is not gauge invariant: if $A_\mu \to A_\mu + \partial_\mu \Lambda$, then $F_{\mu\nu} \to F'_{\mu\nu}$ with

$$F'_{\mu\nu} = F_{\mu\nu} - 2t^{\lambda}_{\mu\nu}\partial_\lambda \Lambda. \tag{3.48}$$

Some authors assert that *torsion does not couple to the gauge field.*

3.3 Kaluza-Klein theory

In 1919 the Polish-German physicist Theodor Kaluza developed a unified theory of the gravitational and electromagnetic forces. His basic idea was to postulate an extra fifth dimension, but with all fields being independent of this extra dimension. The starting point would then be a 5-dimensional pure gravity in which, because of the independence of the fifth coordinate, the fields can be expressesed as 4-dimensional fields.

In 1926 the Swedish physicist Oskar Klein extended this idea. Instead of assuming total independence of the extra dimension, he assumed it to be compact. This means that the fifth dimension has the topology of a circle, with a radius of the order Planck lengh 10^{-35} m. Five dimensional space-time then has the topology $R^4 \times S^1$, and the fifth coordinate y is periodic, $0 \leq my \leq 2\pi$, where m is the inverse radius of the circle. In our normal perception of space-time we would never be able to see this extra dimension.

Let us first define our conventions: hatted quantities will be the five-dimensional ones and the unhatted ones will be the four-dimensional fields. Five-dimensional indices: $\tilde{\mu} = 0, 1, 2, 3, 5$ and of course the four-dimensional indices: $\mu = 0, 1, 2, 3$ ($x^{\tilde{\mu}} = (x^\mu, y)$).

Kaluza generalized the symmetric metric tensor $g_{\mu\nu}$ by adding an additional row and column with the quantities shown as follows:

$$\tilde{g}_{\tilde{\mu}\tilde{\nu}} = \begin{pmatrix} g_{\mu\nu} & kA_\mu \\ kA_\nu & k \end{pmatrix} \tag{3.49}$$

where A_μ is an as-yet undefined vector and k is a constant. Extremalizing the integral form of the line element $ds^2 = \tilde{g}_{\tilde{\mu}\tilde{\nu}}dx^{\tilde{\mu}}dx^{\tilde{\nu}}$ we get the geodesic equation

$$\frac{d^2x^\lambda}{ds^2} + \Gamma^{\lambda}_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = -kF^{\lambda}_{\mu}\frac{dx^\mu}{ds}\frac{dx^5}{ds} - k\tilde{g}^{\lambda 5}\partial_\nu A_\mu \frac{dx^\mu}{ds}\frac{dx^\nu}{ds} \tag{3.50}$$

where $\Gamma^{\lambda}_{\mu\nu}$ are the Christoffel symbols of the second kind and F^{λ}_{μ} is the upper-index form of the Maxwell tensor $F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu$. A similar expression results for the $\tilde{\lambda} = 5$ geodesics. The first term on the right side looks just like the familiar Lorentz force term of a charged particle if we identify $k\dfrac{dx^5}{ds}$ with the charge to mass ratio e/m. But, the second term on the right side contais an additional term involving $\partial_\nu A_\mu$, which is not a tensor quantity. Consequently, Kaluza's $\tilde{g}_{\tilde{\mu}\tilde{\nu}}$ cannot be a true tensor unless we set $\tilde{g}^{\lambda 5} = 0$, which is also problematic.

In 1926, Klein produced the first of two papers that seemed to alleviate that the metric is not a tensor. Klein asserted that the metric actually takes the forms

$$\tilde{g}_{\tilde{\mu}\tilde{\nu}} = \begin{pmatrix} g_{\mu\nu} + kA_\mu A_\nu & kA_\mu \\ kA_\nu & k \end{pmatrix}, \qquad \tilde{g}^{\tilde{\mu}\tilde{\nu}} = \begin{pmatrix} g^{\mu\nu} & -A^\mu \\ -A^\nu & 1/k + A_\mu A^\mu \end{pmatrix} \tag{3.51}$$

which, happily enough, give us the familiar indentity $\tilde{g}_{\tilde{\mu}\tilde{\nu}}\tilde{g}^{\tilde{\mu}\tilde{\lambda}} = \delta^{\lambda}_{\nu}$ and the metric determinant of Klein's metric is $\tilde{g} = kg$. Klein's five-dimensional determinant $\tilde{g}$ is independent of the vector field A_μ.

Klein's four-dimensional geodesic comes out as

$$\frac{d^2x^\lambda}{ds^2} + \Gamma^{\lambda}_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = -kF^{\lambda}_{\mu}\frac{dx^\mu}{ds}\frac{dx^5}{ds} - \frac{1}{2}k\tilde{g}^{\lambda 5}\partial_\nu F^{\lambda}_{\mu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} \tag{3.52}$$

This expression is now fully covariant.

So far we have no real reason to believe that the Kaluza-Klein vector A_μ is related to electromagnetism. In response to this, both Kaluza and Klein considered an infinitesimal change in the fifth coordinate,

$$\begin{aligned} x^5 \to x'^5 &= x^5 + \xi(x^\mu), \quad \text{or} \\ \delta dx'^5 &= dx'^5 - dx^5 = \partial_\mu \xi dx^\mu \end{aligned} \tag{3.53}$$

where ξ is some arbitrary scalar field such that $|\xi| \ll 1$. The five- dimensional line element ds^2, given by

$$ds^2 = \tilde{g}_{\tilde{\mu}\tilde{\nu}}dx^{\tilde{\mu}}dx^{\tilde{\nu}} = g_{\mu\nu}dx^\mu dx^\nu + 2kA_\mu dx^\mu dx^5 + kA_\mu A_\nu dx^\mu dx^\nu + k\left(dx^5\right)^2 \tag{3.54}$$

must be invariant with respect to this variation. The subspace line element $g_{\mu\nu}dx^\mu dx^\nu$ is automatically invariant so we are left with

$$\delta ds^2 = 2kdx^\mu dx^5 \delta A_\mu + 2kA_\mu dx^\mu \delta dx^5 + 2kA_\mu dx^\mu dx^\nu \delta A_\nu + 2kdx^5 \delta dx^5 \tag{3.55}$$

It is a simple matter to show that δds^2 vanishes if only if the variation of the vector A_μ satisfies $\delta A_\mu = -\partial_\mu \xi$: that is,

$$A'_\mu = A_\mu - \partial_\mu \xi \tag{3.56}$$

This is the well-known $U(1)$ gauge transformations of the electromagnetic four-potencial.

The Kaluza-Klein Action

The action approach to gravitation lies in extremalizing the integral

$$I_G = \int \sqrt{-g} R d^4x \tag{3.57}$$

where $R = g^{\mu\nu}R_{\mu\nu}$ is the Ricci scalar. Variation of this integral with respect to the metric gives

$$\delta I_G = \int \sqrt{-g}\left[R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right]\delta g^{\mu\nu}d^4x \tag{3.58}$$

from which we get the Einstein field equation for free space,

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 0 \tag{3.59}$$

Klein naturally assumed that the action would generalize in five dimensions via

$$I_{KK} = \int \sqrt{-\tilde{g}}\tilde{R}d^5x = \sqrt{k}\int \sqrt{g}\tilde{R}d^5x \tag{3.60}$$

Klein calculated all the terms in the five-dimensional Ricci scalar R and he got the simple result

$$\tilde{R} = R + \frac{1}{4}kF_{\mu\nu}F^{\mu\nu} \tag{3.61}$$

where of course $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. He immediatly saw that, without even doing the variational calculation, the five-dimensional action I_{KK} immediately leads to the correct four-dimensional action for the combined gravitational-electromagnetic field:

$$I_{KK} = \int \sqrt{-\tilde{g}}\tilde{R}dx^5 = \sqrt{k}\int \sqrt{-g}\left[R + \frac{1}{4}kF_{\mu\nu}F^{\mu\nu}\right]d^4x\int dx^5 \tag{3.62}$$

Klein was initially bothered by the integral term over dx^5 (which gives infinity) but he quickly recognized that if the fifth dimension was cylindrical, x^5 could be viewed as an angular coordinate having the period $2\pi r$, where r is the cylinder's radius. Klein determined that this radius must be on the order of the Planck constant.

The automatic collapse of Klein's five-dimensional Lagrangian to four dimensions is an example of *dimensional reduction*. This phenomenon has proved to be a powerful tool in modern gauge theories, because a coordinate transformation in the higher space leads to gauge transformation in the subspace.

Chapter 4

Gauge theory and quantum mechanics

4.1 Introduction to quantum mechanics

Very small objects, such as molecules, atoms, and subatomic particles, do not obey the laws of classical mechanics. Quantum mechanics was developed during the 1920's in order to account for their behaviour. The Schrödinger equation is to quantum mechanics what Newton's second law is to classical mechanics: a simple and, in principle, complete statement of the basic physics.

A particle of mass m is described in quantum mechanics by a complex-valued *wave function* $\psi(\mathbf{x}, t)$ which obeys the Schrödinger (non-relativistic) wave equation,

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi \tag{4.1}$$

where $V(\mathbf{x})$ is the same potencial energy function that appeared in classical mechanics, and $h = 2\pi\cdot\hbar$ is a fundamental constant called *Planck's constant*. The standard way to solve such PDE is of course by separation of variables, seeking solutions of the form $\psi(\mathbf{x}, t) = \psi(\mathbf{x})\exp(-iEi/\hbar)$, where E is a constant having the dimensions of energy. This leads to the eigenvalue problem

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{x})\psi = E\psi, \tag{4.2}$$

known as the time-independent Schrödinger equation. The physical interpretation of the wave function, also known as the probability amplitud, is that $|\psi(\mathbf{x}, t)|^2 d^3x$ represents the probability of finding the particle located in an infinitesimal volume element d^3x at point $\mathbf{x}$ at time t. ψ should belong to $L^2(\mathbb{R}^3)$ so that the total probability can be normalized to unity, $\int d^3x|\psi|^2 = 1$, and this provides the boundary condition for the eigenvalue problem. The fundamental tenet of quantum mechanics is that no amount of knowledge concerning the initial conditions or the physical enviroment of the particle will enable one to predict more about its future motion than the probabilities given by ψ. For those accustomed to classical mechanics, this failure of determinism can be difficult to accept.

We will now give the Hamiltonian formulation of quantum mechanics, whose mathematical context is the spectral theory of operators in Hilbert space. The setting is completely different from classical mechanics. The central description of any particular physical system starts with a complex vector space $\mathcal{H}$ associated to that system and called "The Hilbert Space"whether it is one or not (generally it fails to be one only in that the scalar product may not be positive

definite). The spectral properties of the Hermitian operators on $\mathcal{H}$ play a central role. This creates a problem because the operators of interest are generally unbounded and often have no eigenfunctions in $\mathcal{H}$. For example, in $L^2(\mathbb{R}^3)$ we will be interested in the operator $-id/dx$, whose eigenfunctions are $\exp ipx$ with eigenvalue p. Even in the best case, when p is real, these do not belong to $\mathcal{H}$. Even worse, the operator x (that is, multiplication by x) has no eigenfunctions at all unless we allow distributions like $\delta(x-a)$, an "eigenfunction"with eigenvalue a.

In quantum mechanics each state of the physical system is supposed to be represented by a normalized (unit length) vector in $\mathcal{H}$. The state of a system is the totality of information about the system at a given time needed to solve the initial value problem for its evolution starting at that time. In classical mechanics the states were represented by the points of phase space, the values of all coordinates and momenta at a given time. Next, the observables of the system —the measurable quantities such as positions or momenta of particles, their energies or angular momenta and so forth, which were functions on phase space in classical mechanics —are represented by Hermitian (self-adjoint) linear operators (usually unbounded) in $\mathcal{H}$.

The more extended notation in quantum mechanics is the *Dirac notation*: vectors in $\mathcal{H}$ are denoted by $|\psi>$, where ψ is any convenient mnemonic label; e.g. an eigenvector of some operator A with eigenvalue a might be denoted by $|a>$. The scalar product $(\phi, A\psi)$ is written $<\phi, A\psi>$ and is called a matrix element of A; if there is an orthonormal basis of $\mathcal{H}$ containing the vectors $|\phi>$ and $|\psi>$ then this is an entry in the matrix of A relative to this basis. The symbol $<\phi|$ appearing by itself denotes the vector in V^* which is dual to $|\phi>$ via the scalar product.

Suppose then that the system is in state $|\psi>$ and we measure some observable represented by the Hermitian operator A —what result will we obtain? The answer depends on the spectrum of A. It is conventional to discuss separately the cases of discrete and continuous spectrum, although a unified discussion could be given in terms of the spectral projection operators associated to A. We also assume for simplicity that the eigenvalues are all simple (nondegenerate). Thus, let A have eigenvectors $|a>$ obeying $A|a>=a|a>$ and normalized so that $<a'|a>=\delta_{aa'}$ for discrete spectrum, or $<a'|a>=\delta(a-a')$ for continuous spectrum. Then the result of the measurement will be one of the eingenvalues of A, with probability $|<a|\psi>|^2$ for the discrete spectrum. For continuous spectrum $|<a|\psi>|^2 da$ is the probability that the result lies between a and $a+da$. The total probability is 1 because of normalization of $|\psi>$, and the eigenvalues are real because A is Hermitian.

We will frequently use the identity $\sum_a |a><a| = 1$ for the orthonormal eigenvectors of any Hermitician operator, called *inserting a complete set of states*. Applying both sides to any vector gives $\sum_a |a><a|\psi>=|\psi>$, which is just the expansion of $|\psi>$ in a orthonormal basis. It is conventional to write such formulas as if the spectrum were discrete, with the understanding that the sum means an integral over a spectral measure in general. As an example of its use, let us calculate the average value obtained over many measurements of A in state $|\psi>$. This is the sum of the possible values weighted by their probabilities,

$$\sum_a a|<a|\psi>|^2 = \sum_a a<\psi|a><a|\psi>=\sum_a <\psi|Aa><a|\psi>=<\psi|A|\psi>, \tag{4.3}$$

giving a direct physical interpretation to the diagonal matrix elements of A.

Among the observables are the coordinate functions q_i and p_i on phase space themselves, which classically obeyed the Poisson bracket relations

$$[q_i, q_j] = 0, \qquad [p_i, p_j] = 0, \qquad [q_i, p_j] = \delta_{ij}. \tag{4.4}$$

We assume that the corresponding operators in quantum mechanics obey

$$[q_i, q_j] = 0, \qquad [p_i, p_j] = 0, \qquad [q_i, p_j] = i\hbar\delta_{ij}, \tag{4.5}$$

where the bracket is now the commutator, $[A, B] = AB - BA$. The factor i is necessary because the commutator of two Hermitician operators is not Hermitian but skew Hermitian. As a first approximation, any classical function on phase space is assumed to go over to the same function of the operators p and q upon quantization. This prescription is ambiguous because, for example, $pq^2 = q^2p = qpq$ classically but not quantum-mechanically, but in simple physical examples the correct ordering of operators can be fixed.

To complete the postulates of quantum mechanis we must give the rule for time evolution. Among the observables of a system is the total energy; the corresponding Hermitian operator is the Hamiltonian H. The Schrödinger equation states the evolution in time of the states

$$i\hbar\frac{\partial|\psi>}{\partial t} = H|\psi> \tag{4.6}$$

where H is the Hamiltonian operator.

If the Hamiltonian H has no explicit dependence on t and we assume that a system is initially in the state $|\psi(0)>$, then its state at time t will be

$$|\psi(t)>= e^{-iHt/\hbar}\,|\psi(0)>, \tag{4.7}$$

so the time evolution is one group parameter group of unitary transformations generated by H. Any state can be expanded in the basis of eigenstates of H, which satisfy $H|E>= E|E>$ and change only by a phase under time evolution: $|E(t)>= \exp(-iEt)|E(0)>$. Therefore, diagonalizing H solves the time evolution problem for all states. These eigenstates of H are often called *stationary states*, because all probabilities $|<\psi|\ E(t)>|^2$ with fixed $|\psi>$ are constant in time.

Summarizing, quantum mechanics is based on the following axioms:

1. The state of a system is described by a state vector $|\psi>$ (or ψ) in a linear space.
2. The observables (energy, momentum, charge, etc.) are represented by Hermitian operators A, and functions of observables by the corresponding functions of the operators.
3. The expectation value of an observable in the state $<\psi>$ is given by $<A>=<\psi|A|\psi>$.
4. The time evolution is determined by the Schrödinger equation $i\hbar\frac{\partial|\psi>}{\partial t} = H|\psi>$ where H is the Hamiltonian operator.
5. If, in a measurement of the observable A, the value a_n is found, then the original states changes to the corresponding eigenstate $|n>$ of A.

4.1.1 Symmetries in quantum mechanics

When a physical system is invariant under some Lie symmetry group G, we assume that a unitary representation of G acts in the Hilbert space and commutes with the time evolution operator $U_H(t) = \exp(-(i/\hbar)\int H dt)$. For each one-parameter subgroup of G we have an abelian group of unitary linear operator $U(\theta)$ which act on a state of the system $|\psi>$ to give the state of the system after the symmetry transformation, e.g. after rotation by angle θ about some axis.

The unitarity guarantees that scalar products are preserved; since the observable predictions of quantum mechanics are given by various scalar products this ensures that the symmetry transformation does not change the results of experiments. *Stone's theorem* guarantees that each one-parameter subgroup has a Hermitian generator A such that $U(\theta) = \exp(-i\theta A)$, and it follows that the observable A commutes with the Hamiltonian, $[A, H] = 0$. Therefore A is conserved, in the sense that any matrix element

$$< \phi(t)|A|\psi(t) >=< \phi(0)|U_H^\dagger(t)AU_H(t)|\psi(0) >=< \phi(0)|A|\psi(0) > \tag{4.8}$$

is independent of time ($U_H^\dagger(t)$ is the adjoint operator of $U_H(t)$). Thus, in quantum mechanics we have a very direct proof that conserved quantites are the generators of symmetries.

Symmetries are of great practical use in diagonalizing H in concrete problems. Because $[A, H] = 0$, the eigenstates of H can be chosen to simultaneously eigenstates of A, whose spectrum may already be known. In fact, for any element g of the symmetry group G, and any eigenstate $|E >$ of H, it is easy to see that $g|E >$ is another eigenstate having the same eigenvalue. Therefore, each eigenspace of H carries a unitary representation of G! Knowledge of the unitary representations of common symmetry groups such as $SO(3)$ is thus extremely value in understanding the spectrum of an invariant Hamiltonian, which is why it is taught in quantum mechanics courses under titles such as "theory of angular momentum".

4.1.2 Heisenberg picture

Our entire discussion of quantum mechanics so for has been in the so-called *Schrödinger picture* which leads to the Schrödinger equation as the description of dynamics. In this picture the vectors representing the states of the system change with time, but the operators representing physical observables do not. This situation is reversed in the *Heisenberg picture*, which is more useful in Quantum Field Theory.

The Heisenberg picture is obtained by a time-dependent unitary map of $\mathcal{H}$ onto itself. Each state $|\psi >$ at time t is mapped to $U_H(t)|\psi >$, and each operator A is mapped to $U_H^\dagger(t)AU_H(t)$. Because the map is unitary, it preserves all scalar products and matrix elements of operators, hence all physically measurable quantities. Because it inverts the time evolution operator on the states, the state of a system in the Heisenberg picture never changes. Instead, the operators representing a given observable will change with time according $U_H^\dagger(t)AU_H(t)$, or, infinitesimally,

$$\frac{dA}{dt} = -\frac{i}{\hbar}[A, H]. \tag{4.9}$$

This replaces the Schrödinger equation as the description of dynamics. Note that it is identical to the classical Hamilton equation of motion with the Poisson bracket replaced by $-i/\hbar$ times the commutator. This makes it especially clear that the conserved quantities are those which commute with the Hamiltonian. Finally note that the canonical commutation relations $[q_j, p_k] = i\hbar\delta_{ij}$ hold in the Heisenberg only if the operators involved are evaluated at equal times. The commutator $[q_j(t), p_k(t')]$ is more complicated and depends on the specific form of H.

4.1.3 Relativistic quantum mechanics

Efforts to formulate a relativistic quantum mechanics began with attempts to use the correspondence principle in order to derive a relativistic wave equation intended to replace the Schrödinger equation. The first such equation was due to Schrödinger (1926), Gordon (1926), and Klein

(1927). This scalar wave equation of second order, which is now known as the Klein-Gordon equation, was initially dismissed, since it led to negative probability densities.

The year 1928 saw the publication of the Dirac equation. This equation pertains to particles with spin 1/2 and is able to describe many of single-particle properties of fermions. The Dirac equation, like the Klein-Gordon equation, possesses solutions with negative energy, which, in the framework of wave mechanics, leads to difficulties. To prevent transitions of an electron into lower lying states of negative energy should all be occupied. Missing particles in these otherwise occupied states represent particles with opposite charge (antiparticles). The Schrödinger equation, as well as the other axioms of quantum theory, remain unchanged. Only de Hamiltonian is changed.

4.2 The Klein-Gordon equation

4.2.1 Derivation of the Klein-Gordon equation

In order to derive relativistic wave equations, we first recall the *correspondence principle* to replacing the classical quatities by operators. When classical quantities were replaced by the operators

$$\begin{aligned} \text{momentum} \quad & \mathbf{p} \longrightarrow -i\hbar\nabla \\ \text{energy} \quad & E \longrightarrow i\hbar\frac{\partial}{\partial t} \end{aligned} \tag{4.10}$$

we obtained from the non-relativistic energy of a free particle

$$E = \frac{\mathbf{p}^2}{2m} \tag{4.11}$$

the free time-dependent *Schrödinger equation*

$$i\hbar\frac{\partial}{\partial t}\psi = -\frac{\hbar^2\nabla^2}{2m}\psi. \tag{4.12}$$

This equation is obviously not Lorentz covariant due to the different orders of the time and space derivatives.

The relativistic (special) energy-momentum relation is given by

$$E^2 = c^2\mathbf{p}^2 + m^2c^4. \tag{4.13}$$

According to the correspondence principle we obtain

$$-\hbar^2\frac{\partial^2}{\partial t^2}\phi = \left(-\hbar^2c^2\nabla^2 + m^2c^4\right)\phi \tag{4.14}$$

This equation can be written in the even more compact and clearly Lorentz-covariant form

$$\left(\partial_\mu\partial^\mu + \left(\frac{mc}{\hbar}\right)^2\right)\phi = 0 \tag{4.15}$$

Here, $\partial_\mu\partial^\mu = \eta_{\mu\nu}\partial^\mu\partial^\nu \equiv \Box$ is the d'Alambert operator and is invariant under Lorentz transformations

$$\textit{Proof:} \qquad \partial_\mu\eta^{\mu\nu}\partial_\nu = \frac{\partial x'^\lambda}{\partial x^\mu}\frac{\partial}{\partial x'^\lambda}\eta^{\mu\nu}\frac{\partial x'^\rho}{\partial x^\nu}\frac{\partial}{\partial x'^\rho} = \Lambda^\lambda{}_\mu\partial'_\lambda\eta^{\mu\nu}\Lambda^\rho{}_\nu\partial'_\rho = \partial'_\lambda\eta^{\lambda\rho}\partial'_\rho. \tag{4.16}$$

Equation (4.15) was originally introduced and studied by Schrödinger, and by Gordon and Klein. This equation is the same in two reference systems if the field transforms as scalar

$$\phi'\left(x'^{\mu}\right) = \phi\left(x^{\mu}\right) \tag{4.17}$$

Equation (4.15) is known as the free *Klein-Gordon equation* in order to distinguish it from generalizations that additionally contain external potential or electromagnetic fields.

4.2.2 Plane waves

There are two free solutions in the form of plane waves

$$\phi_k = e^{i(Et - \mathbf{p}\cdot\mathbf{x})} \tag{4.18}$$

with

$$E = \pm\sqrt{\mathbf{p}^2c^2 + m^2c^4}.$$

Both positive and negative energies occur here and the energy is not bounded from below.

4.2.3 The continuity equation

To derive a continuity equation we take ϕ^* times (4.15) and substract the complex conjugate of this equation. We obtain the *continuity equation*

$$\dot{\rho} + \operatorname{div}\mathbf{j} = 0 \tag{4.19}$$

with density

$$\rho = \phi^*\frac{\partial\phi}{\partial t} - \phi\frac{\partial\phi^*}{\partial t} \tag{4.20}$$

and current density

$$\mathbf{j} = c^2\left(\phi\nabla\phi^* - \phi^*\nabla\phi\right). \tag{4.21}$$

Here, ρ is not positive definite and thus cannot be directly interpreted as a probability density.

4.2.4 Coupling to an electromagnetic field

The coupling to an electromagnetic field is achieved by making the replacement

$$\partial_\mu \to D_\mu \equiv \partial_\mu + i\frac{e}{c\hbar}A_\mu \tag{4.22}$$

Where D_μ is known as the *covariant derivative*. If we insert (4.22) into the Klein-Gordon equation, we obtain

$$\left\{\left(\partial_\mu + i\frac{e}{c\hbar}A_\mu\right)\left(\partial^\mu + i\frac{e}{c\hbar}A_\mu\right) + \left(\frac{mc}{\hbar}\right)^2\right\}\phi = 0 \tag{4.23}$$

Under gauge transformations ($\alpha(x)$ is an arbitrary function)

$$\begin{aligned} A_\mu(x) &\to A_\mu(x) - \partial_\mu\alpha(x) \\ \phi(x) &\to \exp\left(i\frac{e\alpha(x)}{c\hbar}\right)\phi(x) \end{aligned} \tag{4.24}$$

the equation (4.23) is invariant and D_μ transforms as

$$D_\mu\phi(x) \to \exp\left(i\frac{e\alpha(x)}{c\hbar}\right)D_\mu\phi(x). \tag{4.25}$$

4.3 Dirac equation

4.3.1 Derivation of the Dirac equation

In order that the density be positive, we postulate a differential equation of first order. The requirement of relativistic covariance demands that the spatial derivatives may only be of first order, too. Then, we will now attempt to find a wave equation of the form

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \equiv \left(-i\hbar c\,\alpha^k\partial_k + mc^2\beta\right)\psi \tag{4.26}$$

Spatial components will be denoted by Latin indices, where repeated indices are to be summed over. The Dirac Hamiltonian H is linear in the momentum operator and in the rest energy. We impose the requeriment to the Hamiltonian

$$H^2 = m^2c^4 + c^2\mathbf{p}^2 \tag{4.27}$$

and Lorentz covariance. The resulting equation (4.26) is named, after its discoverer, the *Dirac equation*. We must now look the consequences that arise from the condition (4.27)

$$m^2c^4 - c^2\hbar^2\triangle = m^2c^4\beta^2 - \hbar^2c^2\sum_{ij}\frac{1}{2}\left(\alpha^i\alpha^j + \alpha^j\alpha^i\right)\partial_i\partial_j\psi - i\hbar mc^3\sum_{i=1}^{3}\left(\alpha^i\beta + \beta\alpha^i\right)\partial_i\psi. \tag{4.28}$$

Comparison between left and right sides of (4.28) leads to the three conditions ($i, j = 1, 2, 3$)

$$\alpha^i\alpha^j + \alpha^j\alpha^i = 2\delta^{ij} \tag{4.29}$$

$$\alpha^i\beta + \beta\alpha^i = 0 \tag{4.30}$$

$$\left(\alpha^i\right)^2 = \beta^2 = 1 \tag{4.31}$$

The "objects"α^i and β have to be matrices. From (4.31) it follows that the matrices α^i and β possess only the eigenvalues ± 1.

We may now write (4.30) in the next forms

$$\begin{aligned}\alpha^k &= -\beta\alpha^k\beta\\ \beta &= -\alpha^k\beta\alpha^k,\end{aligned}$$

and using the cyclic invariance of the trace, we obtain

$$\begin{aligned}&\mathrm{Tr}\,\alpha^k = -\mathrm{Tr}\,\beta\alpha^k\beta = -\mathrm{Tr}\,\alpha^k\beta^2 = -\mathrm{Tr}\,\alpha^k \Rightarrow \mathrm{Tr}\,\alpha^k = 0\\ &\mathrm{Tr}\,\beta = -\mathrm{Tr}\,\alpha^k\beta\alpha^k = -\mathrm{Tr}\,\beta(\alpha^k)^2 = -\mathrm{Tr}\,\beta \Rightarrow \mathrm{Tr}\,\beta = 0.\end{aligned} \tag{4.32}$$

Hence, the number of positive and negative eigenvalues must be equal and, therefore, the dimension N of the matrices α^i and β is even. $N = 4$ is the smallest dimension in which it is possible to realize the algebraic structure (4.29), (4.30), (4.31).

We now define new Dirac matrices

$$\begin{aligned}\gamma^0 &\equiv \beta\\ \gamma^i &\equiv \beta\alpha^i\end{aligned}\ . \tag{4.33}$$

These possess the following properties:

$$\begin{aligned}&\left(\gamma^0\right)^\dagger = \gamma^0 \quad \left(\gamma^0\right)^2 = 1\\ &\left(\gamma^i\right)^\dagger = -\gamma^i \quad \left(\gamma^i\right)^2 = -1\end{aligned} \tag{4.34}$$

These relations, together with

$$\gamma^0\gamma^k + \gamma k\gamma^0 = \beta\beta\alpha^k + \beta\alpha^k\beta = 0$$
$$\gamma^k\gamma^l + \gamma l\gamma k = \beta\alpha^k\beta\alpha^l + \beta\alpha^l\beta\alpha^k = 0 \quad \text{for } k \neq l$$

lead to the fundamental algebraic structure of the Dirac matrices

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{4.35}$$

where $\{\quad,\quad\}$ is the anticommutator. The γ^μ matrices have the property (we need that the Hamiltonian be an adjoint operator)

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0 \tag{4.36}$$

The *Dirac equation* in covariant form is written as

$$\left(-i\gamma^\mu\partial_\mu + \frac{mc}{\hbar}\right)\psi = 0. \tag{4.37}$$

ψ is a N-dimensional column vector

$$\psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_N \end{pmatrix} \tag{4.38}$$

and the adjoint to ψ is a N-dimensional row vector

$$\psi^\dagger = (\psi_N^*, \ldots, \psi_N^*)\,. \tag{4.39}$$

And the Dirac adjoint equation is given by

$$i\left(\partial_\mu\psi^\dagger\right)\gamma^{\mu\dagger} + \frac{mc}{\hbar}\psi^\dagger = 0. \tag{4.40}$$

or using (4.36)

$$i\left(\partial_\mu\psi^\dagger\gamma^0\right)\gamma^\mu + \frac{mc}{\hbar}\psi^\dagger\gamma^0 = 0. \tag{4.41}$$

4.3.2 Continuity equation

Multiplying equation (4.37) from the left by $\psi^\dagger\gamma^0$ and adjoint equation (4.41) from right by ψ and add the two resulting equations we have

$$i\partial_\mu\left(\psi^\dagger\gamma^0\gamma^\mu\psi\right) = 0. \tag{4.42}$$

where the four-current is given by

$$j^\mu = \psi^\dagger\gamma^0\gamma^\mu\psi. \tag{4.43}$$

The Dirac equation is invariant under the $U(1)$ action $\psi \to e^{i\theta}\psi$ (global gauge transformation), and the corresponding conserved current is precisely j^μ of Eq.(4.43). The continuity equation in "classical form" is given by

$$\dot{\rho} + \text{div } \mathbf{j} = 0 \tag{4.44}$$

with density

$$\rho = \psi^\dagger\gamma^0\gamma^0\psi = \psi^\dagger\psi = |\psi|^2 \tag{4.45}$$

and current density

$$j^i = c\psi^\dagger\gamma^0\gamma^i\psi \tag{4.46}$$

Here, ρ is positive definite and thus can be directly interpreted as a probabilily density.

4.3.3 Representations of the Dirac matrices

There are infinite representations of the Dirac matrices but *Pauli's fundamental theorem* states the relation between different representations.

Theorem 4.3.1. *For any two four-dimensional representations γ^μ and γ'^μ of Dirac algebra both of which satify the relation* (4.35) *there is a non-singular transformation T such that*

$$\gamma'^\mu = T\gamma^\mu T^{-1}. \tag{4.47}$$

T is uniquely determined to within a constant prefactor.

The more used representations are

Pauli Representation

$$\gamma^0 = \begin{pmatrix} 1_{2x2} & 0 \\ 0 & 1_{2x2} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} 0 & 1_{2x2} \\ 1_{2x2} & 0 \end{pmatrix} \tag{4.48}$$

The matrices σ^i are the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.49}$$

Chiral Representation

$$\gamma^0 = \begin{pmatrix} 0 & -1_{2x2} \\ -1_{2x2} & 0 \end{pmatrix}, \qquad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} 1_{2x2} & 0 \\ 0 & -1_{2x2} \end{pmatrix} \tag{4.50}$$

Majorana Representation

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix}, \quad \gamma^3 = -i\begin{pmatrix} \sigma^1 & 0 \\ 0 & \sigma^1 \end{pmatrix} \tag{4.51}$$

We define the *chirality operator* γ^5 as

$$\gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3. \tag{4.52}$$

Tha matrix γ^5 satisfies the relations

$$\left(\gamma^5\right)^2 = 1 \tag{4.53}$$

and

$$\{\gamma^5, \gamma^\mu\} = 0. \tag{4.54}$$

The next 16 matrices (in 4 dimension) $1_{4x4}, \gamma_5, \gamma_\mu, \gamma_\mu\gamma_5, \frac{i}{2}[\gamma_\mu, \gamma_\nu]$ form a base of the 4×4 matrices and it shows that if Γ_n and $\Gamma_{n'}$ are any of those, then $\mathrm{Tr}\,\Gamma_n\Gamma_{n'} = 0$ if $n \neq n'$.

4.3.4 Lorentz covariance of the Dirac equation

We cosiderer two inertial frames I and I' with the space-time coordinates x and x'. Let the wave function of a particle in these two reference frames be ψ and ψ'. In order that both ψ and ψ' may satisfy the linear Dirac equation, their functional relationship must be linear, i.e,

$$\psi'(x') = S(\Lambda)\psi(x) = S(\Lambda)\psi(\Lambda^{-1}(x'-a)). \tag{4.55}$$

Here, $S(\Lambda)$ is a 4×4 matrix. The Lorentz covariance of the Dirac equation requires the ψ' obey the equation

$$\left(-i\gamma^\mu \frac{\partial}{\partial x'^\mu} + \frac{mc}{\hbar}\right)\psi'(x') = 0 \tag{4.56}$$

The γ matrices are unchanged under the Lorentz transformation. In order to determine S, we need to convert the Dirac equation in the primed and unprimed coordinate systems into one another. The Dirac equation in the unprimed coordinate system is given by (after some manipulations)

$$-iS\Lambda^\nu_\mu \gamma^\mu S^{-1} \frac{\partial}{\partial x'^\nu}\psi'(x') + \frac{mc}{\hbar}\psi'(x') = 0 \tag{4.57}$$

From a comparison of (4.57) with (4.56), it follows that the Dirac equation is form invariant under Lorentz transformations, provided $S(\Lambda)$ satisfies the following condition:

$$S(\Lambda)^{-1}\gamma^\nu S(\Lambda) = \Lambda^\nu_\mu \gamma^\mu \tag{4.58}$$

A wave function that transforms under a Lorentz transformation according to $\psi' = S\psi$ is known as a *four-component Lorentz spinor.*

In order to know the transformation behaviour of the most important bilinear quantities under Lorentz transformations we need to establish a relationship between the adjoint transformation $S^\dagger$ and S^{-1}.

Theorem 4.3.2.

$$S^\dagger \gamma^0 = b\gamma^0 S^{-1}, \tag{4.59}$$

where

$$b = \pm 1 \quad for \quad \Lambda^{00} \begin{cases} \geqslant +1 \\ \leqslant -1 \end{cases}$$

Proof: (for example, in ref[10]).

Theorem 4.3.3. *Under orthochronous Lorentz transformations $\psi^\dagger(x)\gamma^0\psi(x)$ transforms as a scalar.*

$$\textit{Proof:} \qquad \psi^\dagger(x')\gamma^0\psi(x') = \psi^\dagger(x)S^\dagger\gamma^0 S\psi(x) = \psi^\dagger\gamma^0 S^{-1}S\psi = \psi^\dagger\gamma^0\psi. \tag{4.60}$$

Theorem 4.3.4. *Under orthochronous Lorentz transformations the four-current behaves as a vector.*

$$\begin{aligned} \textit{Proof:} \qquad j'^\mu(x') &= \psi'^\dagger(x')\gamma^0\gamma^\mu\psi(x') = \psi^\dagger(x)S^\dagger\gamma^0\gamma^\mu S\psi(x) \\ &= \psi^\dagger(x)\gamma^0 S^{-1}\gamma^\mu S\psi(x) = \Lambda^\mu_\nu \psi^\dagger(x)\gamma^0\gamma^\nu\psi(x) = \Lambda^\mu_\nu j^\nu(x). \end{aligned} \tag{4.61}$$

In particular, for a finite rotacions (about the z axis) and boost (along the x^i direction) we have

$$\begin{aligned} S_{rot} &= \cos\frac{\theta}{2} + i\sigma^{12}\sin\frac{\theta}{2} \\ S_{boost} &= \cosh\frac{\xi}{2} + \gamma^0\gamma^i \sinh\frac{\xi}{2} \end{aligned} \tag{4.62}$$

where

$$\sigma^{12} = \frac{i}{2}[\gamma^1, \gamma^2] \qquad \cosh\xi = \frac{1}{\sqrt{1 - \mathbf{v}^2/c^2}} \qquad \sinh\xi = \frac{|\mathbf{v}|}{c}\frac{1}{\sqrt{1 - \mathbf{v}^2/c^2}} \tag{4.63}$$

The transformation S for rotations is unitary ($S^{-1} = S^\dagger$). We can see that

$$S_{rot}(2\pi) = -1 \qquad S_{rot}(4\pi) = 1. \tag{4.64}$$

This means that spinors do not regain their initial value after a rotation through 2π, but only a rotation through 4π.

4.3.5 Solutions of the Dirac equation for free particles

We seek solutions of the free Dirac equation (4.37) in the form of plane waves:

$$\psi(x) = \frac{1}{(2\pi\hbar)^{3/2}} e^{-ipx/\hbar} u\left(p\right) \tag{4.65}$$

$u\left(p\right)$ is known as *Dirac spinor* and satisfies the algebraic equation

$$\left(\gamma^{\mu} p_{\mu} - mc\right) u(p) = 0 \tag{4.66}$$

Solutions in the form

$$\psi(x) = \frac{1}{(2\pi\hbar)^{3/2}} e^{+ipx/\hbar} v\left(p\right) \tag{4.67}$$

where $v\left(p\right)$ satisfies the equation

$$\left(\gamma^{\mu} p_{\mu} + mc\right) v(p) = 0 \tag{4.68}$$

are possible too. In the rest frame of the particle $p^{\mu} = (mc, \mathbf{0})$ and we have

$$\begin{array}{l} \left(\gamma^0 - 1\right) u(m, \mathbf{0}) = 0 \\ \left(\gamma^0 + 1\right) v(m, \mathbf{0}) = 0. \end{array} \tag{4.69}$$

In the Pauli representation we have the solutions (for particles at rest):

$$u^1(m,\mathbf{0}) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \ , \ u^2(m,\mathbf{0}) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \ , \ v^1(m,\mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \ , \ v^2(m,\mathbf{0}) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{4.70}$$

and for any reference frame:

$$u^r(p) = \frac{\gamma^{\mu} p_{\mu} + mc}{\sqrt{2m(mc^2 + E)}} u^r(m,\mathbf{0}) = \begin{pmatrix} \left(\dfrac{E + mc^2}{2mc^2}\right)^{1/2} \chi_r \\ \\ \dfrac{\sigma \cdot \mathbf{p}}{\left(2m(mc^2 + E)\right)^{1/2}} \chi_r \end{pmatrix} \tag{4.71}$$

$$v^r(p) = \frac{-\gamma^{\mu} p_{\mu} + mc}{\sqrt{2m(mc^2 + E)}} v^r(m,\mathbf{0}) = \begin{pmatrix} \dfrac{\sigma \cdot \mathbf{p}}{\left(2m(mc^2 + E)\right)^{1/2}} \chi_r \\ \\ \left(\dfrac{E + mc^2}{2mc^2}\right)^{1/2} \chi_r \end{pmatrix} \tag{4.72}$$

where $\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

The most general spinor can be represeted by the following superposition of the positive and the negative energy states

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{m}{E} \sum_{r=1,2} \left(b(p,r) u^r(p) e^{-ip_{\mu}x^{\mu}} + d(p,r) u^r(p) e^{-ip_{\mu}x^{\mu}} \right) \tag{4.73}$$

The normalization

$$\int d^3x \, \psi^{\dagger}(t,\mathbf{x}) \psi(t,\mathbf{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{m}{E} \sum_{r=1,2} \left(|b(p,r)|^2 + |d(p,r)|^2 \right) = 1 \tag{4.74}$$

is time independent as a result of the continuity equation.

4.3.6 Coupling to an electromagnetic field

The coupling to the electromagnetic field is achieved by making the replacement (4.22). If we insert (4.22) into the Dirac equation equation, we obtain

$$\left(-i\gamma^\mu\left(\partial_\mu + \frac{e}{c\hbar}A_\mu\right) + \left(\frac{mc}{\hbar}\right)\right)\psi = 0 \tag{4.75}$$

If $\psi(x)$ satisfies the Dirac equation for the potencial A_μ, the transformed spinor $\psi'(x') = S\psi(x)$ satisfies the Dirac equation for the potencial $A'_\mu(x') = \Lambda_{\mu\nu}A^\nu(x)$.

Under gauge transformations ($\alpha(x)$ is an arbitrary function)

$$\begin{aligned} A_\mu(x) &\to A_\mu(x) - \partial_\mu\alpha(x) \\ \psi(x) &\to \exp\left(i\frac{e\alpha(x)}{c\hbar}\right)\psi(x) \end{aligned} \tag{4.76}$$

the equation (4.75) is invariant and D_μ transforms as

$$D_\mu\psi(x) \to \exp\left(i\frac{e\alpha(x)}{c\hbar}\right)D_\mu\psi(x). \tag{4.77}$$

4.3.7 Discrete symmetries

In addition to the behaviour of the spinor under continuous Lorentz symmetry, we must consider their behaviour under discrete symmetries, such as generated by parity, charge conjugation, time-reversal symmetry, and chirality.

PARITY

The transformation corresponding to a spatial reflection ($\mathbf{x} \to$ **-x**) is called *parity*. The Lorentz transformation corresponding to a spatial reflection is represented by

$$\Lambda^\mu_{\ \nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{4.78}$$

The associated S is determined , according to (4.58), from

$$S^{-1}\gamma^\mu S = \Lambda^\mu_{\ \nu}\gamma^\nu = \eta^{\mu\mu}\gamma^\mu \qquad \text{(no sum)}, \tag{4.79}$$

An explicit solution of this equation, which we shall denote in this case by $\mathcal{P}$, is simply given by:

$$S = \mathcal{P} \equiv e^{i\varphi}\gamma^0. \tag{4.80}$$

where $e^{i\varphi}$ is an irrelevant phase factor.

Thus, the action of the parity transformation on a spinor field is given by:

$$\psi'(-\mathbf{x}, t) = \mathcal{P}\psi(\mathbf{x}, t). \tag{4.81}$$

We also have $\mathcal{P}^\dagger = \mathcal{P}$ and $\mathcal{P}^2$. If A^μ is invariant under parity, then the Dirac Hamiltonian H satisfies

$$[\mathcal{P}, H] = 0. \tag{4.82}$$

CHARGE CONJUGATION

Charge conjugation is easily studied by taking the Dirac equation and then reversing sign of the electric charge. If we let ψ_c represent the Dirac field that has the opposite charge to ψ, then we have

$$\begin{aligned}\left(-i\gamma^\mu\left(\partial_\mu+\frac{e}{c\hbar}A_\mu\right)+\left(\frac{mc}{\hbar}\right)\right)\psi=0\\ \left(-i\gamma^\mu\left(\partial_\mu-\frac{e}{c\hbar}A_\mu\right)+\left(\frac{mc}{\hbar}\right)\right)\psi_c=0\end{aligned}\tag{4.83}$$

In order to find the relationship between ψ with charge e and ψ_c, with charge $-e$, let us take the complex conjugate and then the transpose of the first equation. Then we find:

$$(\gamma^\mu)^\top\left(-i\partial_\mu-\frac{e}{c\hbar}A_\mu\right)\left(\left(\gamma^0\right)^\top\psi^*\right)=0\tag{4.84}$$

It can be shown that for any representation of the Dirac algebra, there exist a matrix C that satisfies

$$C\gamma_\mu^T C^{-1}=-\gamma_\mu\tag{4.85}$$

Now let us compare the previous equation with the equation for the ψ_c field. We have an exact correspondence (up to a phase) if we set

$$\psi_c=e^{i\varphi}C\left(\left(\gamma^0\right)^\top\psi^*\right)\tag{4.86}$$

So far, we have not specified the representation of the Dirac matrices. there is more than one solution of the equation for C. In the Dirac representation, we find the following solution for the C matrix as:

$$C=i\gamma^2\gamma^0\tag{4.87}$$

which satisfies the following additional constraints

$$C=-C^{-1}=C^T=C^\dagger\tag{4.88}$$

The important feature of the C matrix is that it allows us to identify the particle-antiparticle structure of the Dirac equation. We can proof easily that the four-current j^μ transforms under charge conjugation operation changing the sign.

TIME REVERSAL

Time reversal transformation is the transformation $t\to t'$. We can write down the Dirac equation with a time reversal and try to retransform the equation back into the usual Dirac form. We find the result

$$\mathcal{T}\psi(\mathbf{x},t)\mathcal{T}^{-1}=e^{i\varphi}T\psi(\mathbf{x},-t),\tag{4.89}$$

where

$$T\gamma^\mu T^{-1}=\gamma_\mu^\top=\gamma^{\mu*}\tag{4.90}$$

An explicit representation of the T matrix is given by:

$$T=i\gamma^1\gamma^3\tag{4.91}$$

where:

$$T=T^\dagger=T^{-1}=-T^*\tag{4.92}$$

We should also mention that the $\mathcal{T}$ operator is inusual because it is antiunitary. Correspondigly we wish that $\mathcal{T}$ would reverse the exponent appearing in the time evolution operator:

$$\mathcal{T}e^{iH(t_1-t_2)}\mathcal{T}^{-1}=e^{iH(t_2-t_1)}\tag{4.93}$$

However, this is impossible if the Hamiltonian commutes with $\mathcal{T}$.

CHIRALITY

The chiral projection operators are defined by

$$P_{\pm} = \frac{1}{2}(1 \pm \gamma_5). \tag{4.94}$$

They satisfy

$$P_{\pm}^2 = P_{\pm}, \quad [P_+, P_-] = 0 \tag{4.95}$$

Using these projection operators one can decompose every fermion field into chiral components

$$\psi = \psi_L + \psi_R \tag{4.96}$$

with

$$\psi_L \equiv P_+\psi, \quad \psi_R \equiv P_-\psi. \tag{4.97}$$

The operators $P_{\pm}$ are called *chiral* projection operators because in certain limits they project out particles of a certain helicity (chirality= "handedness"). The *helicity* operator, which measures the spin along the direction of motion of the particle takes the form

$$h(\mathbf{p}) = \frac{\hbar}{2|\mathbf{p}|}\gamma_5\gamma_0\gamma\mathbf{p}, \tag{4.98}$$

where $\mathbf{p}$ is the momentum of the particle. For massless particles,

$$h(\mathbf{p})\psi_L(x) = -\frac{\hbar}{2}\psi_L(x), \quad h(\mathbf{p})\psi_R(x) = +\frac{\hbar}{2}\psi_R(x) \tag{4.99}$$

Hence, for massless fermions (neutrinos) the eigenstates of $P_{\pm}$ coincide with the helicity eigenstates. Experimentaly, it is found that only neutrinos of negative chirality (left-handed) exist and their antiparticle, the antineutrinos, have positive quirality.

4.4 Pauli's dimensional reduction

In 1953 Pauli made the following statement which shows his clarividency:

" *I am very much in favour of the general principle to bring empirical conservation laws and invariance properties in connection with mathematical groups of transformations of Nature. If besides the conservation of energy-momentum and of charge the conservation of the property defined as the number of nucleons and charge-independence of the nuclear forces are well-established they have indeed, as Pais tried now to express mathematically, also to be connected with group theoretical properties of the laws of nature... I would like to ask in this conection whether the transformation group (isospin group) with constant phases can be amplified in a way analogous to the gauge-group for electromagnetic potentials in such a way that the meson-nucleon interaction is connected with the amplified group. The main problem seems to be the incorporation of the coupling constant into the group.* "

Pauli knew Weyl's 1929 paper and the Kaluza-Klein theory and he presented his findings in two letters that he wrote to Pais in the same year 1953. The idea was that since the original Kaluza reduction produced the mathemathical structure for describing electromagnetism, a generalization of that theory to more dimensions might produce the mathematical structure for describing the strong interactions. The letters contain the first correct expression for the

non-abelian field strengths. Pauli did not send his notes for publication. Probably he knew the problem of the mass (in the 1953 physicists thought that the bosons π were the carriers of the strong nuclear interactions and these had mass).

Pauli's theory, like that of Kaluza and Klein, was based on dimensional reduction, but it differed from the Kaluza-Klein theory in two major respects. First, Pauli increased the number of extra dimensions from 1 to 2, thereby introducing non-abelian groups for non-gravitational interactions for the first time. Second, he identified the electromagnetic potentials with components of the Christoffel connection, whereas Kaluza and Klein had identied them with components of the metric tensor.

We realize a generalization of Pauli's theory and we consider a $(4+n)$-dimensional space with coordinates $x_A = (x_\mu, y_a)$ where the y's are the internal coordinates and $y = 0$ is space-time. Let us restrict the possible coordinate transformations to coordinate transformations of space-time and linear transformations of the internal coordinates which may depend on the space-times

$$y^a \to R^a_b(x) y^b \quad a, b = 1 \dots n. \tag{4.100}$$

Let us now consider the parallel transport of a vector $v(x)$ which depends only on the space-time coordinates, along a curve is space-time

$$\begin{aligned} (\nabla v_\mu)^\alpha &= \partial_\mu v^\alpha + \Gamma^\alpha_{\mu\beta_0} v^\beta + \Gamma^\alpha_{\mu b_0} v^b \\ (\nabla v_\mu)^a &= \partial_\mu v^a + \Gamma^a_{\mu\beta_0} v^\beta + \Gamma^a_{\mu b_0} v^b \end{aligned} \tag{4.101}$$

where ∇ and Γ are the the usual Riemannian covariant derivative and the Christoffel symbols for the full space respectively and the subscript zero means evaluation at $y = 0$. Suppose that

$$\Gamma^\alpha_{\mu b_0} = \Gamma^a_{\mu\beta_0} = 0 \tag{4.102}$$

then

$$\begin{aligned} (\nabla v_\mu)^\alpha &= \partial_\mu v^\alpha + \Gamma^\alpha_{\mu\beta_0} v^\beta \\ (\nabla v_\mu)^a &= \partial_\mu v^a + \Gamma^a_{\mu m_0} v^b \end{aligned} \tag{4.103}$$

which means that the space-time components v^μ and the internal-space components v^a transform independently. Thus, under the assumption (4.102), the restriction of the parallel transfer of the full space to space-time decomposes into the ordinary parallel transfer in space-time for the space-time vectors v^ν and a parallel transfer of the form

$$D_\mu v^a = \partial_\mu v^a + A^a_{\mu b} v^b, \tag{4.104}$$

where

$$A^a_{\mu b} v^b = \Gamma^a_{\mu\beta_0} v^\beta, \tag{4.105}$$

for the Kaluza-space vectors $v^a(x)$. The connection A_μ transforms as a vector with respect to coordinate transformations in space-time, but with respect to the coordinate-gauge transformations (4.100), it transforms according to

$$A^a_{\mu b} \to R^a_c R^d_b A^c_{\mu d} + R^c_a \partial_\mu R^c_b \tag{4.106}$$

This follows from the general law for the metric components,

$$g_{ab} \to R^c_a R^d_b A^c_{\mu d} \quad \text{and} \quad g^a_\mu \to R^a_b g^b_\mu + \partial_\mu R^a_c y^c, \tag{4.107}$$

and the definition

$$A^a_{\mu b} = \frac{1}{2} g^{ac} \left(\frac{\partial g_{\mu b}}{\partial y} - \frac{\partial g_{\mu c}}{\partial y_b} - \frac{\partial g_{bc}}{\partial x_\mu} \right)_0 . \tag{4.108}$$

We can see that $A^a_{\mu b}$ is a gauge-connection from the point of view of space-time. And,

$$R^a_{b\mu\nu} = (\partial_\mu A_\nu - \partial_\nu A_\mu)^a_b + A^a_{\mu c}A^c_{\nu b} - A^a_{\nu c}A^c_{\mu b} = (\partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu])^a_b \,. \tag{4.109}$$

$R^a_{b\mu\nu}$ is part of the Riemann tensor for the full space but is also the Riemann tensor for the gauge connection, $A_\mu(x)$.

4.5 Non-abelian gauge fields

Let us generally considerer fields that transform according to some representation of a certain Lie group G. This means that, for every element of the group G, we have a matrix U; these matrices U satisfy the same multiplication rules as the corresponding elements of G. Under a group transformation the fields rotate as follows

$$\psi(x) \to \psi'(x) = U\psi(x) \tag{4.110}$$

For most groups the matrices U can be written in exponencial form

$$U = \exp\left(\xi^a t_a\right) \tag{4.111}$$

where the matrices t_a are the generators of the group defined in the representation appropiate to ψ, and the ξ^a constitute a set of linearly independent *real* parameters in terms of which the group elements can be described. Let us considerer the extension of the group G to a group of *local* gauge transformations. This means that the parameters of G will become functions of the space-time coordinates x^μ and the effect of local transformation on derivatives of the fields,

$$\begin{aligned} &\psi(x) \to \psi'(x) = U(x)\psi(x) \\ &\partial_\mu\psi(x) \to (\partial_\mu\psi(x))' = U(x)\partial_\mu\psi(x) + (\partial_\mu U(x))\,\psi(x). \end{aligned} \tag{4.112}$$

Due to the presence of the second term on the right-hand side, ∂_μ does not transform covariantly. Therefore one attemps to replace ∂_μ by a so-called *covariant derivative* D_μ, which constitutes a covariant quantity when applied to ψ,

$$D_\mu\psi(x) \to (D_\mu\psi(x))' = U(x)D_\mu\psi(x). \tag{4.113}$$

A covariant derivative can be viewed as the result of a particular combination of the ordinary derivative and a field gauge transformation,

$$D_\mu\psi \equiv \partial_\mu\psi - W_\mu\psi \tag{4.114}$$

where W_μ is a matrix of the type generated by an infinitesimal gauge transformation. These means that W_μ takes values in the Lie-algebra corresponding to the group G, i.e., W_μ can be descomposed into the generators t_a,

$$W_\mu = W^a_\mu t_a. \tag{4.115}$$

and must transforms under gauge transformations as

$$(W_\mu\psi)' = (\partial_\mu\psi)' - (D_\mu\psi)' = \left\{UW_\mu U^{-1} + (\partial_\mu U)\,U^{-1}\right\}\psi' \tag{4.116}$$

This implies the following transformation rule for W_μ,

$$W_\mu \to UW_\mu U^{-1} + (\partial_\mu U)\,U^{-1}. \tag{4.117}$$

Clearly the gauge fields do not transform covariantly. Before verifying the consistency of this result we note that the transformation rule for W_μ defines a group, because successive application of two transformations U_1 and U_2 on W_μ gives the same result as directly applying a single transformation $U_3 = U_2 U_1$,

$$\begin{aligned} W_\mu \to W''_\mu &= U_2(U_1 W_\mu U_1^{-1} + (\partial_\mu U_1)\, U_1^{-1})U_2^{-1} + (\partial_\mu U_2)U_2^{-1} \\ &= (U_2U_1)W_\mu(U_2U_1)^{-1} + \partial_\mu(U_2U_1)(U_2U_1)^{-1}. \end{aligned} \tag{4.118}$$

The consistence of (4.118) requires that the right-hand side are also Lie-algebra valued and is not difficult verify that.

Unlike ordinary differentiations, two covariant differentations do not necessarily commute. The commutator of two covariant derivatives D_μ and D_ν is given by

$$[D_\mu, D_\nu]\,\psi = D_\mu\,(D_\nu\psi) - D_\nu\,(D_\mu\psi) = -\,(\partial_\mu W_\nu - \partial_\nu W_\mu - [W_\mu, W_\nu])\,\psi. \tag{4.119}$$

This results leads to the definition of a covariant antisymmetric tensor $G_{\mu\nu}$,

$$G_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - [W_\mu, W_\nu]\,, \tag{4.120}$$

which is called the *field strength*. As ψ and $D_\mu D_\nu \psi$ transform identically under the gauge transformations the field strength must transform covariantly according to

$$G_{\mu\nu} \to G'_{\mu\nu} = U G_{\mu\nu} U^{-1}. \tag{4.121}$$

Because W_μ is Lie-algebra valued and the quadratic term in (4.120) is a commutator, the field strength is also Lie-algebra valued, i.e., $G_{\mu\nu}$ can also be decomposed in terms of the group generators t_a,

$$G_{\mu\nu} = G^a_{\mu\nu} t_a, \tag{4.122}$$

with

$$G^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - f^a_{bc} W^b_\mu W^c_\nu. \tag{4.123}$$

The result (4.124) can now be expressed in a representation independent form,

$$[D_\mu, D_\nu] = -G_{\mu\nu}. \tag{4.124}$$

This equation is known as the *Ricci identity*. We may apply further covariant derivatives to (4.124). In particular, the *Bianchi identity*

$$D_\mu G_{\nu\rho} + D_\nu G_{\rho\mu} + D_\rho G_{\mu\nu} = 0. \tag{4.125}$$

where, according to (4.121), the covariant derivative of $G_{\mu\nu}$ equals

$$D_\mu G_{\nu\rho} = \partial_\mu G_{\nu\rho} - [W_\mu, G_{\nu\rho}], \tag{4.126}$$

or, in components,

$$D_\mu G^a_{\nu\rho} = \partial_\mu G^a_{\nu\rho} - f^a_{bc} W^b_\mu G^c_{\nu\rho}. \tag{4.127}$$

Under infinitesimal gauge transformations ψ, W_μ and $G_{\mu\nu}$ transform as

$$\begin{aligned} &\psi \to \psi + \xi^a t_a \psi, \\ &W^a_{\mu\nu} \to W^a_{\mu\nu} - \partial_\mu \xi^a - f^a_{bc} W^b_\mu \xi^c = W^a_{\mu\nu} + D_\mu \xi^a, \\ &G_{\mu\nu} \to G_{\mu\nu} + [\xi^a t_a, G_{\mu\nu}]\,. \end{aligned} \tag{4.128}$$

4.5.1 Lagrangian formulation

Let (M, g) be a space-time manifold of dimension $d = 1+3$ endowed with a pseudo Riemannian structure. We have the Hodge operator $* : \Omega^2(M) \to \Omega^2(M)$, which obeys $* \circ * = -1$.

Let P be a principal G bundle over M, where G is a compact simple Lie group (not necessarily simply connected).

We wish to define an action functional S on the space of connections on P. The main requeriment is that S should be gauge invariant, i.e. invariant under the infinite dimensional Lie group Aut (P) of bundle automorphisms of P (which has an induced action on the space of connections.) The action S should also be local, i.e. it should be given as

$$S = \int_M \mathcal{L} \tag{4.129}$$

where the Lagrangian density $\mathcal{L}$ is a polynomial in the fields. In pure Yang-Mills theory, the only available field is the curvature $F \in \Omega^2(M, \mathrm{ad}(P))$. The most general Lagrangian of at most second order in F is then

$$\mathcal{L} = \frac{1}{2g^2}\mathrm{Tr}\,(F \wedge *F) + \frac{\theta}{8\pi^2}\mathrm{Tr}\,(F \wedge F) \quad (+ \text{ constant}). \tag{4.130}$$

Here Tr is an invariant bilinear form on the Lie algebra G of G, and g and θ are known as the coupling constant and the theta angle respectively.

We now wish to determine the conditions on a connection for the action S to be stationary with respect to an arbitrary small perturbation $\delta A \in \Omega^1(M, \mathrm{ad}(P))$. Locally, i.e. over an open subset U of M, we can describe a connection on P by a gauge potential $A \in \Omega^1(U, \mathrm{ad}(as))$. Actually, we need to impose some condition like δA being compactly supported or sufficiently rapidly decaying so that the partial integration may be performed on M without generating any surface term on ∂M. One then finds that the Euler-Lagrange equations are given by the Yang-Mills equations

$$d(*F) = 0. \tag{4.131}$$

Together with the Bianchi identity

$$dF = 0 \tag{4.132}$$

they are a non-abelian generalization of Maxwell's equations (without sources). Note that the coupling constant g, nor the theta angle θ appears in these equations, so they play no role in the classical theory.

4.6 The paper of Yang and Mills

4.6.1 Introduction

The construction of a non-abelian gauge theory was made by Chen-Ning Yang and Robert L. Mills. Motivated by the desire to make the isotopic spin symmetry of the strong interactions local, they invented the $SU(2)$ gauge theory. Yang had been considering the problem since 1949. Yang was impressed with the idea that charge conservation was related to the invariance of the theory under phase changes and he tried to generalize this to isotopic spin interactions. And they found that they were unable to conclude what the mass of the gauge-particles should be. The question of the gauge-field mass problem was raised by Pauli when Yang was invited to present the Yang-Mills results at the Princeton Institute (1954).

PHYSICAL REVIEW VOLUME 96, NUMBER 1 OCTOBER 1, 1954

Conservation of Isotopic Spin and Isotopic Gauge Invariance[1]

C.N. YANG AND R. L. MILLS

Brookhaven National Laboratory, Upton, New York

(Received June 28, 1954)

It is pointed out that the usual principle of invariance under isotopic spin rotation is not consistant with the concept of localized fields. The possibility is explored of having invariance under local isotopic spin rotations. This leads to formulating a principle of isotopic gauge invariance and the existence of a **b** field which has the same relation to the isotopic spin that the electromagnetic field has to the electric charge. The **b** field satisfies non-linear differential equations. The quanta of the **b** field are particles with spin unity, isotopic spin unity, and electric charge $\pm e$ or zero.

INTRODUCTION

THE conservation of isotopic spin is a much discussed concept in recent years. Historically an isotopic spin parameter was introduced by Heisenberg[2] two describe the two charge states (namely neutron and proton) of a nucleon. The idea that the neutron and proton correspond to two states of the same particle was suggested at the time by the fact that their masses are nearly equal, and that the light stable even nuclei contain equal numbers of them. Then in 1937 Breit, Condon, and Present pointed out the approximately equality of $p-p$ and $n-p$ interactions in the S state.[3] It seemed natural to assume that this equality holds also in the other states available to both the $n-p$ and $p-p$ systems. Under such an assumption one arrives at the concept of a total isotopic spin[4] which is conserved in nucleon-nucleon interactions. Experiments in recent years[5] on the energy levels of light nuclei strongly suggest that this assumption is indeed correct. An implication of this this is that all strong interactions such as the pion-nucleon interaction, must also satisfy the same conservation law. This and the knowledge that there are three charges states of the pion, and that pions can be coupled to the nucleon field *singly*, lead to the conclusion that pions have isotopic spin unity. A direct verification of this conclusion was found in the experiment of Hildebrand[6] which compares the differential cross section of the process $n+p \rightarrow \pi^0 + d$ with that of the previously measured process $p+p \rightarrow \pi^+ + d$.

The conservation of isotopic spin is identical with the requirement of invariance of all interactions under isotopic spin rotation. This means that when electromagnetic interactions can be neglected, as we shall hereafter assume to be the case, the orientation of the isotopic spin is of no physical significance. The differentation between a neutron and a proton is then a purely arbitrary process. As usually conceived, however, this arbitrariness is subject to the following limatiation: once one choose what to call a proton, what a neutron, at one space-time point, one is the not free to make any choices at other space-time points.

It seems that this is not consistent with the localized field concept that underlies the usual physical theories. In the present paper we wish to explore the possibility of requiring all interactions to be

[1]Work performed under the auspices of the U. S. Atomic Energy Commission.

[2]W. Heisenberg, Z. Physik **77**, **1** (1932).

[3]Breit, Condon, and Present, Phys. Rev. **50**, 825 (1936). J. Schwinger pointed out that the small difference may be attributed to magnetic interactions [Phys. Rev. **78**, 135 (1950)].

[4]The total isotopic spin **T** was first introduced by E. Wigner, Phys. Rev. **51** 106 (1937); B. Cassen and E. U. Condon, Phys. rev. **50**, 846 (1936).

[5]T. Lauritsen, Ann. Rev. Nuclear Sci. **1**, 67 (1952); D. R. Inglis, Revs. Modern Phys. **25**, 390 (1953).

[6]R. H. Hildebrand, Phys. Rev. **89**, 1090 (1953).

invariant under *rotations* of the isotopic spin at all space-time points bocomes a physically meaningless quantity (the electromagnetic field being neglected).

We wish to point out that an entirely similar situation arises with respect to the ordinary gauge invariance of a charged field which is described by a complex wave function ψ. A change of gauge[7] means a change factor $\psi \to \psi'$, $\psi' = (\exp i\alpha)\psi$, a change that is devoid of any physical consequences. Since ψ may depend on x, y, z, and t, the relative phase factor of ψ at two different space-time points is therefore completely arbitrary. In other words, the arbitrariness in choosing the phase factor is local in character.

We define *isotopic gauge* as an arbitrary way of choosing the orientation of the isotopic spin axes at all space-time points, in analogy with the electromagnetic gauge represents an arbitrary way of choosing the complex phase factor of a charged field at all space-time points. We then propose that all physical processes (not involving the electromagnetic field) be invariant under an isotopic gauge transformation, $\psi \to \psi'$, $\psi' = S^{-1}\psi$, where S represents a space-time dependent isotopic spin rotation.

To preserve invariance one notices that in electrodynamics it is necessary to counteract the variation of α with x, y, z, and t by introducing the electromagnetic field A_μ which changes under a gauge transformation as

$$A'_\mu = A_\mu + \frac{1}{e}\frac{\partial \alpha}{\partial x_\nu}.$$

In an entirely manner we introduce a B field in the case of the isotopic gauge transformation to counteract the dependence of S on x, y, z, and t. It will be seen that this natural generalization allows for very little arbitrariness. The field equations satisfied by the twelve independent components of the B field, which we shall call **b** field, and their interaction with any field having an isotopic spin are essentially fixed, in much the same way that the free electromagnetic field and its interaction with charged fields are essentially determined by the requirement of gauge invariance.

In the following two sections we put down the mathematical formulation of the idea of isotopic gauge invariance discussed above. We then proceed to the quantization of the field equations for the **b** field. In the last section the properties of the quanta of the **b** field are discussed.

ISOTOPIC GAUGE TRANSFORMATION

Let ψ be a two-component wave function describing a field with isotopic spin $\frac{1}{2}$ (or arbitrary else). Under an isotopic gauge transformation it transforms by

$$\psi = S\psi', \tag{1}$$

where S is a 2×2 unitary with determinant unity. In accordance with discussion in the previous section, we require, in analogy with electromagnetic case, that all derivatives of ψ appear in the following combination:

$$(\partial_\mu - i\epsilon B_\mu)\,\psi.$$

B_μ are 2×2 matrices such that[8] for $\mu = 1, 2$, and 3, B_μ is Hermitian and B_4 is anti-Hermitian. Invariance requires that

$$S\left(\partial_\mu - i\epsilon B'_\mu\right)\psi' = (\partial_\mu - i\epsilon B_\mu)\,\psi. \tag{2}$$

Combining (1) and (2), we obtain the isotopic gauge transformation on B_μ:

$$B'_\mu = S^{-1}B_\mu S + \frac{i}{\epsilon}S^{-1}\frac{\partial S}{\partial x_\mu}. \tag{3}$$

The last term is similar to the gradient term in the gauge transformation of electromagnetic potencials. In analogy to the procedure of obtaining gauge invariant field strengths in the electromagnetic case, we define now

$$F_{\mu\nu} = \frac{\partial B_\mu}{\partial x_\nu} - \frac{\partial B_\nu}{\partial x_\mu} + i\epsilon(B_\mu B_\nu - B_\nu B)_\mu). \tag{4}$$

[7] W. Pauli, Revs. Modern Phys. **13**, 203 (1941).

[8] We use the conventions $\hbar = c = 1$, and $x_4 = it$. Bold-face typre refers to vectors in isotopic space, not in space-time.

One easily shows from (3) that

$$F'_{\mu\nu} = S^{-1} F_{\mu\nu} S \tag{5}$$

under an isotopic gauge transformation.[9] Other simple functions of B than (4) do not lead to such a simple transformation property.

The above lines of thought can be applied to any field ψ with arbitrary isotopic spin. One need only use other representations S of rotations in three dimensional space. It is reasonable to assume that different fields with the same total isotopic spin, hence belonging to the same representation S, interact with the same matrix field B_μ. (This is analogous to the fact that the electromagnetic fields interacts in the same way with any charged particle, regardless of the nature of the particle. If different fields interact with different and independent B fields, there would be more conservation laws than simply the conservation of total isotopic spin.) To find a more explicit form for the B fields and to relate the B_μ's corresponding to different representations S, we proceed as follows.

Equation (3) is valid for any S and its corresponding B_μ. Now the matrix $S^{-1}\partial S/\partial x_\mu$ appearing in (3) is a linear combination of the isotopic spin "angular momentum"matrices T^i $(i = 1, 2, 3)$ corresponding to the isotopic spin of the ψ field we are considering. So B_μ itself must also contain a linear combination of the matrices T^i. But any part of B_μ in addition to this, $\bar{B}_\nu$, say, is a scalar or tensor combination of the T's, and must transform by the homogeneous part of (3), $\bar{B}'_\mu = S^{-1}\bar{B}_\mu S$. Such a field is extraneous; it was allowed by the very general form we assumed for the B field, but is irrelevant to the question of isotopic gauge. Thus the relevant part of the B field is of the form

$$B_\mu = 2\mathbf{b}_\mu \cdot \mathbf{T}. \tag{6}$$

(Bold-face letters denote three-component vectors in isotopic space.) To relate the $\mathbf{b}_\mu$'s corresponding to different representations S we now consider the product representation $S = S^{(a)}S^{(b)}$. The B field for the combination transform, according to (3), by

$$B'_\mu = [\,S^b\,]^{-1}[\,S^a\,]^{-1} B S^{(a)} S^{(b)} + \frac{i}{\epsilon}\left[S^{(a)}\right]^{-1}\frac{\partial S^{(a)}}{\partial x_\mu} + \frac{i}{\epsilon}\left[S^{(b)}\right]^{-1}\frac{\partial S^{(b)}}{\partial x_\mu}.$$

But the sum of $B^{(a)}_\mu$ and $B^{(b)}_\mu$, the B fields corresponding to $S^{(a)}$ and $S^{(b)}$, tranforms in exactly the same way, so that

$$B_\mu = B^{(a)}_\mu + B^{(b)}_\mu$$

(plus possible terms which transform homogeneously, and hence are irrelevant and will not be included). Decomposing $S^{(a)}S^{(b)}$ into irreducible representations, we see that twelve-component field $\mathbf{b}_\mu$, in Eq. (6) is the same for all representations.

To obtain the interaction between any field ψ of arbitrary isotopic spin with the $\mathbf{b}$ one therefore simply replaces the gradient of ψ by

$$(\partial_\mu - 2i\epsilon\mathbf{b}_\mu \cdot \mathbf{T})\,\psi, \tag{7}$$

where T^i $(i = 1, 2, 3)$, as defined above, are the isotopic spin "algular momentum "matrices for the field ψ.

We remark that the nine components of $\mathbf{b}_\mu$, $\mu = 1, 2, 3$ are real ans the three of $\mathbf{b}_4$ are pure imaginary. The isotopic-gauge covariant field quantities $F_{\mu\nu}$ are expressible in terms of $\mathbf{b}_\mu$:

$$F_{\mu\nu} = 2\mathbf{f}_{\mu\nu} \cdot \mathbf{T}, \tag{8}$$

where

$$\mathbf{f}_{\mu\nu} = \partial_\nu \mathbf{b}_\mu - \partial_\mu \mathbf{b}_\nu - 2\epsilon\mathbf{b}_\mu \times \mathbf{b}_\nu. \tag{9}$$

$\mathbf{f}_{\mu\nu}$ transforms like a vector under an isotopic gauge transformation. Obviously the same $\mathbf{f}_{\mu\nu}$ interact with all fields ψ irrespective of the representation S that ψ belongs to.

[9] *Note added in proof.* It may appear that B_μ could be introduced as an auxiliary quantity to accomplish invariance, but need not be regarded as a field variable by itself. It is to be emphasized that such a procedure violates the principle of invariance. Every quantity that is not a pure numeral (like 2, or M, or any definite representation of the γ matrices) should be regarded as a dynamical variable, and should be varied in the Lagrangian to yield an equation of motion. Thus the quantities B_μ must be regarded as independent fields.

The corresping transformation of $\mathbf{b}_\mu$ is cumbersome. One need, however, study only the infinitesimal gauge tranformations,

$$S = 1 - 2i\mathbf{T} \cdot \delta\omega.$$

Then

$$\mathbf{b}'_\mu = \mathbf{b}_\mu + 2\mathbf{b}_\mu \times \delta\omega + \frac{1}{\epsilon}\frac{\partial}{\partial x_\mu}\delta\omega. \tag{10}$$

FIELD EQUATIONS

To write down the field equations for the $\mathbf{b}$ field we clearly only want to use isotopic gauge invariant quantities. In analogy with the electromagnetic case we therefore write down the following Lagrangian density: [10]

$$\mathcal{L} = -\frac{1}{4}\mathbf{f}_{\mu\nu} \cdot \mathbf{f}_{\mu\nu}$$

Since the inclusion of a field with isotopic spin $\frac{1}{2}$ is illustrative, and does not complicate matters very much, we shall use the following total Lagrangian density:

$$\mathcal{L} = \bar{\psi}\gamma_\mu(\partial_\mu - i\epsilon\tau \cdot \mathbf{b}_\mu)\psi - m\bar{\psi}\psi. \tag{11}$$

One obtains from this the following equations of motion:

$$\begin{aligned} &\frac{\partial \mathbf{f}_{\mu\nu}}{\partial x_\nu} + 2\epsilon(\mathbf{b}_\nu \times \mathbf{f}_{\mu\nu}) + \mathbf{J}_\mu = 0, \\ &\gamma_\mu(\partial_\mu - i\epsilon\tau \cdot \mathbf{b}_\mu)\psi + m\psi = 0, \end{aligned} \tag{12}$$

where

$$\mathbf{J}_\mu = i\epsilon\bar{\psi}\gamma_\mu\tau\psi. \tag{13}$$

The divergence of $\mathbf{J}_\mu$ does not vanish. Instead it can easily be shown from (13) that

$$\frac{\partial J_\mu}{\partial x_\mu} = -2\epsilon\mathbf{b}_\mu \times \mathbf{J}_\mu. \tag{14}$$

If we define, however,

$$\mathfrak{J}_\mu = \mathbf{J}_\mu + 2\epsilon\mathbf{b}_\nu \times \mathbf{f}_{\mu\nu}, \tag{15}$$

Then (12) lead to the equation of continuity,

$$\partial\mathfrak{J}_\mu/\partial x_\mu = 0. \tag{16}$$

$\mathfrak{J}_{1,2,3}$ and 4 are respectively the isotopic spin current density and isotopic spin density of the system. The equation of continuity guarantees that the total isotopic spin

$$\mathbf{T} = \int \mathfrak{J}_4 d^3x$$

is independent of time and independent of a Lorents transformation. It is important to notice that $\mathfrak{J}_\mu$, like $\mathbf{b}_\mu$, does not transform exactly like vectors under isotopic space rotations. But the total isotopic spin,

$$\mathbf{T} = -\int \frac{\partial \mathbf{f}_{4i}}{\partial x_i} d^3x$$

is the integral of the divergence of $\mathbf{f}_{4i}$, which transforms like a true vector under isotopic spin space rotations. Hence, under a general isotopic gauge transformation, if $S \to S_0$ on an infinitely large sphere, $\mathbf{T}$ would transform like an isotopic spin vector.

[10] Repeated indices are summed over, except where explicitly stated otherwise. Latin indices are summed from 1 to 3, Greek ones from 1 to 4.

Equation (15) shows that the isotopic spin arises both from the spin-$\frac{1}{2}$ field ($\mathbf{J}_\mu$) and from the $\mathbf{b}_\mu$, field itself. Inasmuch as the isotopic spin is the source of the $\mathbf{b}$ field, this fact makes the fiel equations for the $\mathbf{b}$ field non-linear, even in the absence of the spin-$\frac{1}{2}$ field. This is different from the case of the electromagnetic field, which is itself chargeless, and consequently satisfies equations in the absence of a charged field.

The Hamiltonian derived from (11) is easily demonstrated to be positive definite in the absence of the field of isotopic spin $\frac{1}{2}$. The demostration is completely identical with the similar one in electrodynamics. We must complete the set of equations of motion (12) and (13) by the supplementary condition,

$$\partial \mathbf{b}_\mu / \partial x_\mu = 0, \tag{17}$$

which serves to eliminate the scalar part of the field in $\mathbf{b}_\mu$. This clearly imposes a condition on the possible isotopic gauge transformations. that is, the infinitesimal isotopic gauge transformation $S = 1 - i\tau \cdot \delta\omega$ must satisfy the following condition:

$$2\mathbf{b}_\mu \times \frac{\partial}{\partial x_\mu}\delta\omega + \frac{1}{\epsilon}\frac{\partial^2}{\partial x_\mu^2}\delta\omega = 0. \tag{18}$$

This is the analog of the equation $\partial^2\alpha/\partial x_\mu^2 = 0$ that must be satisfied by the gauge transformation $A'_\mu = A_\mu + e^{-1}(\partial\alpha(\partial x_\mu)$ of the electromagnetic field.

QUANTIZATION

To quantize, it is not convenient to use the isotopic gauge invariant Lagrangian density (11). This is quite similar to the corresponding situation in electrodynamics and we adopt the customary procedure of using a Lagrangian density which is not obviously gauge invariant:

$$\mathcal{L} = -\frac{1}{2}\frac{\partial \mathbf{b}_\mu}{\partial x_\nu} \cdot \frac{\partial \mathbf{b}_\mu}{\partial x_\nu} + 2\epsilon(\mathbf{b}_\mu \times \mathbf{b}_\nu)\frac{\partial \mathbf{b}_\mu}{\partial x_\nu} - \epsilon^2(\mathbf{b}_\mu \times \mathbf{b}_\nu)^2 + \mathbf{J}_\mu \cdot \mathbf{b}_\mu - \bar{\psi}(\gamma_\mu\partial_\mu + m)\psi. \tag{19}$$

The equations of motion that result from this Lagrangian density can be easily shown to imply that

$$\frac{\partial^2}{\partial x_\nu^2}\mathbf{a} + 2\epsilon\mathbf{b}_\nu \times \frac{\partial}{\partial x_\nu}\mathbf{a} = 0,$$

where

$$\mathbf{a} = \partial \mathbf{b}_\mu / \partial x_\mu.$$

Thus if, consistent with (17), we put on one space-like surface $\mathbf{a} = 0$ together with $\partial\mathbf{a}/\partial t = 0$, it follows that $\mathbf{a} = 0$ at all times. Using this supplementary condition one can easily prove that the field equations resulting from the Lagrangian densities (19) and (11) are identical. One can follow the canonical method of quantization with the Langrangian density (19). Defining

$$\mathbf{\Pi}_\mu = -\frac{\partial \mathbf{b}_\mu}{\partial x_4} + 2\epsilon(\mathbf{b}_\mu \times \mathbf{b}_4),$$

one obtains the equal-time commutation rule

$$[b_\mu^i(x), \Pi_\nu^j(x')]_{t-t'} = -\delta_{ij}\delta_{\mu\nu}\delta^3(x - x'), \tag{20}$$

where b_μ^i, $i = 1, 2, 3$, are the three components of $\mathbf{b}_\mu$. The relativistic invariance of these commutation rules follows from the general proof for canonical methods of quantization given by Heisenberg and Pauli.[11]

[11] W. Heisenberg and W. Pauli, Z. Physik **56**, 1 (1929).

The Hamiltonian derived from (19) is identical with the one from (11), in virtue of the supplementary condition. Its density is

$$H = H_0 + H_{int},$$

$$H_0 = -\frac{1}{2}\mathbf{\Pi}_\mu \cdot \mathbf{\Pi}_\mu + \frac{1}{2}\frac{\partial \mathbf{b}_\mu}{\partial x_j}\frac{\partial \mathbf{b}_\mu}{\partial x_j} + \bar{\psi}(\gamma_j \partial_j + m)\psi, \tag{21}$$

$$H_{int} = 2\epsilon(\mathbf{b}_i \times \mathbf{b}_4) \cdot \mathbf{\Pi}_i - 2\epsilon(\mathbf{b}_\mu \times \mathbf{b}_j) \cdot (\partial \mathbf{b}_\mu / \partial x_j) + \epsilon^2(\mathbf{b}_i \times \mathbf{b}_j)^2 - \mathbf{J}_\mu \cdot \mathbf{b}_\mu.$$

The quantized form of the supplementary condition is the same as in quantum electrodynamics.

PROPERTIES OF THE b QUANTA

The quanta of the **b** field clearly have spin unity and isotopic spin unity. We know theire electric charge too because all the interactions that we proposed must satisfy the law of conservation of electric charge, which is exact. The two states of the nucleon, namely proton and neutron, differ by charge unity. Since they can transform into each other through the emission or absorption of a **b** quantum, the latter must have three charge states with charges $\pm e$ and 0. Any measurement of electric charges of course involves the electromagnetic field, which necessarily introduces a preferential direction in isotopic space at all space-time points. Choosing the isotopic space, one sees that for the nucleons

$$Q = \text{electric charge} = e(\frac{1}{2} + \epsilon^{-1}T^z),$$

and for the **b** quanta

$$Q = (e/\epsilon)T^z.$$

The interaction (7) then fixes the electric charge up to an additive constant for all the fields with any isotopic spin:

$$Q = e(\epsilon^{-1}T^z + R). \tag{22}$$

The constants R for two charge conjugate fields must be equal but have opposite signs. [12]

We next com to the question of the mass of the **b** quantum, to which we do not have a satisfactory answer. One may argue that without a nucleon field the Lagrangian would contain no quantity of the dimension of mass, and that therefore the mass of the **b** quantum in such case is zero. This argument is however subject to the criticism that, like all field theories, the **b** field is beset with divergences, and dimensional arguments are not satisfactory.

One may of course try to apply to the **b** field the methods for handling infinities development for quantum electrodynamics. Dyson's approach[13] is best suited for the present case. One first transforms into the interaction representation in which the state vector Ψ satisfies

$$i\partial\Psi/\partial t = H_{\text{int}}\Psi,$$

where H_{int} was defined in Eq.(21). The matrix elements of the scattering matrix are then formulated in terms of contributions from Feynman diagrams. These diagrams have three elementary types of vertices illustrated in Fig. 1, instead of only one type as in quantum electrodynamics.

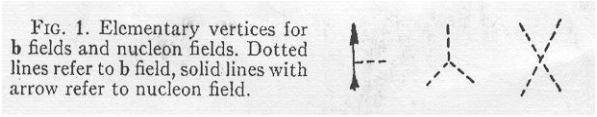


FIG. 1. Elementary vertices for b fields and nucleon fields. Dotted lines refer to b field, solid lines with arrow refer to nucleon field.

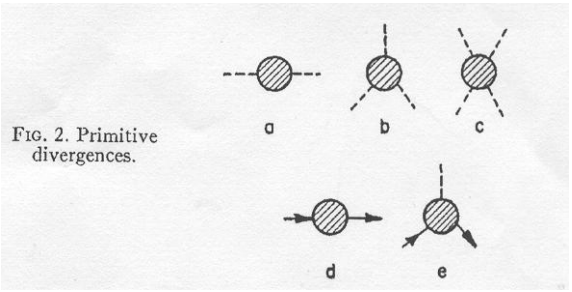


FIG. 2. Primitive divergences.

[12] See M. Gell-Mann, Phys. Rev. **92**, 833 (1953).
[13] F. J. Dyson, Phys. Rev. **75**, 486, 1736 (1949).

The "primitive divergences"are still finite in number and are listed in Fig. 2. Of these, the one labeled a is the one that effects the propagation function of the **b** quantum, and whose singularity determines the mass of the **b** quantum. In electrodynamics, by the requirement of electric charge conservation,[14] it is argued that the mass of the photon vanishes. Corresponding arguments in the **b** field case do no exist[15] even though the conservation of isotopic spin still holds. We have therefore not been able to conclude anything about the mass of the **b** quantum.

A conclusion about the mass of the **b** quantum is of course very important in deciding whether the proposal of the existence of the **b** field is consistent with experimental information. For example, it is inconsistent with present experiments to have their mass less tan the pions, because among other reasons they would then be created abundantly at high energies and the charged ones should live long enough to be seen. If they have a mass greater than that of pions, on the other hand, they would have a short lifetime (say, less than 10^{-20} sec) for decay into pions and photons and would so far have escaped detection.

4.6.2 Comments to the paper

Yang and Mills developped a non-abelian gauge field theory based on $SU(2)$ for strong interactions. Now, we present this theory in a modern form based in the theory of Lie groups.

Consider a set of N spinor fields ψ_i transforming under transformations U belonging to a certain group G according to $(i, j = 1, \ldots, N)$. Writting the matrices in exponentiated form,

$$U(\xi) = \exp(g\xi t_a), \quad (a = 1, 2, 3) \tag{4.133}$$

where g is the coupling constant and t_a are the generators of $SU(2)$. The three generators of $SU(2)$ are expressed in terms of the isotopic spin matrices τ_a,

$$t_a = \frac{1}{2} i\tau_a, \tag{4.134}$$

which coincide with the Pauli matrices used in the context of ordinary spin. Consequently the generators t_a satisfy the commutation relations

$$[t_a, t_b] = -\epsilon_{abc} t_c. \tag{4.135}$$

Identifing terms

$$\epsilon B_\mu = \frac{1}{2} igW_\mu^a \tau_a, \tag{4.136}$$

where

$$b_\mu^a = W_\mu^a \quad (a = 1, \ldots, 3) \tag{4.137}$$

And the field equations can be written as:

$$\begin{array}{l} \partial_\nu G_{\mu\nu}^a - g\varepsilon_{abc} W^b G_{\mu\nu}^c + J^a = 0 \\ \gamma^\mu(\partial_\mu - igB_\mu)\psi + m\psi = 0 \end{array} \tag{4.138}$$

In terms of covariant derivatives we have

$$D_\mu G_{\mu\nu}^a = -J^a \tag{4.139}$$

[14] J. Schwinger, Phys. rev. **76**, 790 (1949).

[15] In electrodynamics one can formally prove the $G_{\mu\nu} k_\nu = 0$, where $G_{\mu\nu}$ is defined by Schwinger's Eq. (A12). ($G_{\mu\nu} A_\nu$ is the current generated through virtual processes by the arbitrary external field A_ν.) No corresponding proof has been found for the present case. This is due to the fact that in electrodynamics the conservation of charge is a consequence of the equation of motion of electron field alone, quite independently of the electromagnetic field itself. In the present casethe **field** carries an isotopic spin and destroys such general conservation laws.

In turns out that, contrary to initial expectations, local $SU(2)$ transformations have no role to play in the strong interactions. Instead these forces are governed by an $SU(3)$ gauge theory called *quantum chromodynamics* because one has introduced the term *colour* for the degrees of freedom transforming under $SU(3)$. This theory will be discussed in chapter 7.

Chapter 5

Canonical quantization of fields

5.1 The real Klein-Gordon field

The quantization of mechanics follows from the Heisenberg commutation relations

$$\begin{aligned} &[x_i, p_j] = i\delta_{ij} \quad (i,j = 1,2,3) \\ &[x_i, x_j] = [p_i, p_j] = 0, \end{aligned} \tag{5.1}$$

where the momentum p_i is defined canonically as $\partial L/\partial \dot{x}_i$; $\mathbf{x}$ and $\mathbf{p}$ refer to the position and momentum of the particle, measured at the same time. In a scalar field theory, $\phi(\mathbf{x}, t)$ plays a role analogous to $\mathbf{x}(t)$, and describes a system with an infinite number of degrees of freedom, since, at each time, ϕ has an independent value at each point in space. To approach this continuum case, let us divide space up into cells, each of volume δV_r, and let $\phi_r(t)$ be the average value of $\phi(x)$ in cell r at time t. Let the average lagrangian in each cell be $\mathcal{L}_r$. Then the momentum variable p_r, conjugate to ϕ_r is

$$p_r(t) = \frac{\partial L}{\partial \dot{\phi}_r(t)} = \delta V_r \frac{\partial \mathcal{L}_r}{\partial \dot{\phi}_r(t)} = \delta V_r \pi_r(t) \tag{5.2}$$

where the field $\pi(\mathbf{x}, t)$ is defined by

$$\pi(\mathbf{x}, t) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\mathbf{x}, t)} \tag{5.3}$$

and $\pi_r(t)$ is its average value in cell r. Then the Heisenberg commutation relations give

$$\begin{aligned} &[\phi_r(t), p_s(t)] = i\delta_{rs}, \\ &[\phi_r(t), \phi_s(t)] = [p_r(t), p_s(t)] = 0. \end{aligned} \tag{5.4}$$

Substituting (5.2) into (5.4) gives $[\phi_r(t), \pi_s(t)] = (1/\delta V_s) i\delta_{rs}$. In the continuum limit $\delta V_r \to 0$, and we have

$$\begin{aligned} &[\phi(\mathbf{x}, t), \pi(\mathbf{x}', t)] = i\delta(\mathbf{x} - \mathbf{x}'), \\ &[\phi(\mathbf{x}, t), \phi(\mathbf{x}', t)] = [\pi(\mathbf{x}, t), \pi(\mathbf{x}', t)] = 0. \end{aligned} \tag{5.5}$$

These are known as *equal-time commutation relations.*

The scalar field ϕ obeys the Klein-Gordon (KG) equation

$$\left(\Box + m^2\right) \phi = 0, \tag{5.6}$$

The KG equation can be derived from the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu \phi \partial^\mu \phi\right) - \frac{1}{2} m^2 \phi^2. \tag{5.7}$$

The Fourier expansions for the Klein-Gordon field, ϕ may be written

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3\, 2\omega_k} \left[a(k)e^{-ikx} + a^\dagger(k)e^{ikx}\right] \tag{5.8}$$

whith $\omega_k = \sqrt{\mathbf{k}^2 + m^2}$ ($\hbar = c = 1$). The coefficients $a(k)$ and $a^\dagger(k)$ are also operators. The measure in the integrand has been so chosen because it is relativistically invariant. We can obtain the next commutation relations

$$\begin{aligned} \left[a(k), a^\dagger(k')\right] &= (2\pi)^3\, 2\omega_k \delta^3(\mathbf{k} - \mathbf{k}'), \\ [a(k), a(k')] &= \left[a^\dagger(k), a^\dagger(k')\right] = 0. \end{aligned} \tag{5.9}$$

We construct the operator

$$N(k) = a^\dagger(k) a(k). \tag{5.10}$$

It is simple to show that $N(k)$ and $N(k')$ commute, so the eigenstates of these operators may be used to form a basis. Let the eigenvalues be denoted by $n(k)$:

$$N(k)|n(k)\rangle = n(k)|n(k)\rangle \tag{5.11}$$

Then

$$\begin{aligned} N(k)a^\dagger(k)|n(k)\rangle &= (n(k)+1)\, a^\dagger(k)|n(k)\rangle \\ N(k)a(k)|n(k)\rangle &= (n(k)-1)\, a(k)|n(k)\rangle. \end{aligned} \tag{5.12}$$

These equation tells us that if the state $|n(k)\rangle$ has eigenvalue $n(k)$, the states $a^\dagger|n(k)\rangle$ and $a(k)|n(k)\rangle$ are eigenstates of $N(k)$ with respective eigenvalues $n(k)+1$ and $n(k)-1$. $N(k)$ is the *particle number operator*. The operators $a^\dagger$ and $a(k)$ are called the *creation* and *annihilation* operators, respectevily. The operators Hamiltonian and momentum are

$$H = \int \frac{d^3k}{(2\pi)^3 2} \left(N(k) + \frac{1}{2}\right) \tag{5.13}$$

and

$$\mathbf{P} = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \mathbf{k} \left(N(k) + \frac{1}{2}\right). \tag{5.14}$$

We can combine H and $\mathbf{P}$ into a 4-momentum operator (where $k^0 = \omega_k$)

$$P^\mu = \int \frac{d^3k}{(2\pi)^3 2\omega_k} k^\mu \left(N(k) + \frac{1}{2}\right), \tag{5.15}$$

which is indeed the generator of the translation group in $\mathbb{R}^4$,

$$[\phi(x), P^\mu] = i\partial^\mu \phi(x). \tag{5.16}$$

To see that $N(k)$ never becomes negative note that the state $a(k)|n(k)\rangle$ must have non-negative norm:

$$\left[a(k)|n(k)\rangle\right]^\dagger \left[a(k)|n(k)\rangle\right] = \langle n(k)|a^\dagger(k)a(k)|n(k)\rangle = n(k)\langle n(k)|n(k)\rangle > 0; \tag{5.17}$$

so that, if $|n(k)\rangle$ has non-negative norm, $n(k)$ must be positive or zero. On the other hand, $a(k)$ reduces $n(k)$ by 1, and repeated application will continue to reduce it. The only way to avoid $n(k)$ becoming negative is to have a ground state (the vacuum) $|O(k)\rangle$, or $|0\rangle$ for short, with

$$a(k)|0\rangle = 0 \tag{5.18}$$

and so

$$N(k)|0\rangle = a^\dagger(k)a(k)|0\rangle = 0; \tag{5.19}$$

the vacuum contains no particles with momentum $\mathbf{k}$. Application of $a^\dagger$ now increases $N(k)$ in steps of 1 at a time, so $N(k)$ is *integral*.

It is not difficult to show that an arbitrary, normalised state containing $n(k_1)$ particles with momentum k_1, $n(k_2)$ with momentum k_2, etc., may be written

$$|n(k_1), n(k_2), \ldots\rangle = \frac{1}{\sqrt{(n(k_1)!n(k_2)!\ldots)}} \left[a^\dagger(k_1)\right]^{n(k_1)} \left[a^\dagger(k_2)\right]^{n(k_2)} \cdots |0\rangle. \tag{5.20}$$

There is, evidently no restriction on $n(k)$; any number of particles may exist in the same momentum state. These particles are therefore *bosons*. These particles are the quanta of the field, and this is the very deep connection between fields and particles in Quantum Field Theory: the stationary states of a free field theory are multiparticle states. The Klein-Gordon field describes spinless particles such as π mesons.

5.2 The complex Klein-Gordon field

If the scalar field now has two components ϕ_1 and ϕ_2, we may put

$$\phi = (\phi_1 + i\phi_2)/\sqrt{2}, \qquad \phi^* = (\phi_1 - i\phi_2)/\sqrt{2} \tag{5.21}$$

The fields (which are regarded independent) ϕ and ϕ^* obey the KG equations

$$\left(\Box + m^2\right)\phi = 0, \tag{5.22}$$

$$\left(\Box + m^2\right)\phi^* = 0. \tag{5.23}$$

These equations can be derived from the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) - \frac{1}{2}m^2\phi^*\phi \tag{5.24}$$

The Lagrangian is clearly invariant under the gauge global transformation

$$\begin{array}{l} \phi \;\; \rightarrow \mathrm{e}^{-i\Lambda}\phi \\ \phi^* \rightarrow \mathrm{e}^{-i\Lambda}\phi^* \end{array} \tag{5.25}$$

where Λ is a real constant. The group concerned is $U(1)$.

The concerned current given by Noether's theorem is

$$j^\mu = i\left(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*\right) \tag{5.26}$$

and we will have the corresponding conserved quantity Q (the electric charge).

The Fourier expansions for the complex (or charged) Klein-Gordon field are given by

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3\, 2\omega_k} \left[a(k)e^{-ikx} + b^\dagger(k)e^{ikx}\right], \tag{5.27}$$

$$\phi^*(x) = \int \frac{d^3k}{(2\pi)^3\, 2\omega_k} \left[b(k)e^{-ikx} + a^\dagger(k)e^{ikx}\right]. \tag{5.28}$$

and the equal-time commutators different from zero are

$$\begin{aligned} \left[a(k), a^\dagger(k')\right] &= (2\pi)^3\, 2\omega_k \delta^3(\mathbf{k}-\mathbf{k}'), \\ \left[b(k), b^\dagger(k')\right] &= (2\pi)^3\, 2\omega_k \delta^3(\mathbf{k}-\mathbf{k}'). \end{aligned} \tag{5.29}$$

The Hamiltonian becomes

$$H = \int \frac{d^3k}{(2\pi)^3 2} \Big(a^\dagger(k)a(k) + b^\dagger(k)b(k)\Big). \tag{5.30}$$

$a^\dagger$ and $b^\dagger$ can be interpreted as creation operators for particles and antiparticles, which carry opposite charge, but have the same mass.

5.3 The Dirac field

The Lagrangian density of the Dirac field is given by

$$\mathcal{L} = \frac{i}{2}\left[\bar\psi\gamma^\mu(\partial_\mu\psi) - (\partial_\mu\bar\psi)\gamma^\mu\psi\right] - m\bar\psi\psi. \tag{5.31}$$

In this Lagrangian, ψ and $\bar\psi$ are treated as dynamically independent fields. The canonical momentum field is

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot\psi(x)} = i\psi^\dagger(x). \tag{5.32}$$

The Hamiltonian density is then

$$\mathcal{H} = \pi\dot\psi - \mathcal{L} = \psi^\dagger\gamma^0\left(-i\gamma^i\partial_i + m\right)\psi = \psi^\dagger\gamma^0(i\gamma^0\partial_0\psi) = \psi^\dagger i\frac{\partial\psi}{\partial t} \tag{5.33}$$

The general solution to the Dirac equation may expanded in terms of the plane wave solutions as follows:

$$\psi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{m}{\omega_k}\sum_{\alpha=1,2}\left[b_\alpha(k)u^\alpha(k)e^{-ikx} + d^\dagger_\alpha(k)v^\alpha(k)e^{ikx}\right]$$

$$\bar\psi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{m}{\omega_k}\sum_{\alpha=1,2}\left[b^\dagger_\alpha(k)\bar u^\alpha(k)e^{ikx} + d_\alpha(k)\bar v^\alpha(k)e^{-ikx}\right]$$

where $u^{(1,2)}$ and $v^{(1,2)}$ are the positive and negative energy spinors. And an annihilation operator $b_\alpha(k)$ multiplies the positive energy term, and a creation operator $d^\dagger(k)$ the negative energy term. Substituting (6.0) and (6.1) into (6.0) and using the normalisation conditions for the spinors u and v gives for the energy

$$H = \int d^3x\mathcal{H} = \int \frac{d^3k}{(2\pi)^3} m\sum_{\alpha=1,2}\left[b^\dagger_\alpha(k)b_\alpha(k) - d_\alpha(k)d^\dagger_\alpha(k)\right].$$

The only way to avoid the negative energy terms is to introduce the *anticommutators*, defined by

$$\{A, B\} = AB + BA$$

and to postulate the anticommutation relations

$$\{b_\alpha(k), b^\dagger_{\alpha'}(k')\} = \{d_\alpha(k), d^\dagger_{\alpha'}(k')\} = (2\pi)^3 \frac{\omega_k}{m} \delta^3(\mathbf{k} - \mathbf{k}')\delta_{\alpha\alpha'},$$

$$\{b_\alpha(k), b_{\alpha'}(k')\} = \{b^\dagger_\alpha(k), b^\dagger_{\alpha'}(k')\} = 0,$$

$$\{d_\alpha(k), d_{\alpha'}(k')\} = \{d^\dagger_\alpha(k), d^\dagger_{\alpha'}(k')\} = 0.$$

To substract out the zero point energy, we normal order the Hamiltonian, with the additional introduction, in the case of Fermi fields, to change the sign of the term for each interchange of operators. This gives

$$H = \int d^3x : \psi^\dagger i \frac{\partial \psi(x)}{\partial t}(x) := \int \frac{d^3k}{(2\pi)^3} m \sum_{\alpha=1,2} \left[b^\dagger_\alpha(k) b_\alpha(k) + d_\alpha(k) d^\dagger_\alpha(k) \right].$$

This is now positive definite, but it is easy to see that the anticommutation relations imply Fermi statistics; for example $\{b^\dagger_\alpha(k), b^\dagger_\alpha(k)\} = 0$ implies $b^\dagger_\alpha(k) b^\dagger_\alpha(k) = 0$, hence $b^\dagger_\alpha(k) b^\dagger_\alpha(k)|0\rangle = 0$; it is impossible to have two quanta of the Dirac field in the same state. Hence the use of anticommutators leads directly to the Pauli exclusion principle.

5.4 The electromagnetic field

We turn now to the gauge fields, and consider, for simplicity, the electromagnetic field (the other fundamental gauge fields in the *Standard Model* are the the weak and gluon fields). There are complications met in quantizing gauge fields. The origin of difficulties is that the electromagnetic field, like any massless field, possesses only two independent components, but is covariantly described by a 4-vector A_μ.

The Maxwell's equations in a vacuum follow from Lagrangian density

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{5.34}$$

with

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{5.35}$$

For a given electromagnetic field $F_{\mu\nu}$, A_μ is not unique; the gauge transformation $A_\mu \to A'_\mu = A_\mu + \partial_\mu \Lambda(x)$ leaves $F_{\mu\nu}$ unchanged. By choosing Λ to satisfy

$$\Box \Lambda = -\partial_\mu A^\mu, \tag{5.36}$$

we obtain $\partial_\mu A'^\mu = 0$.. Dropping the prime, the resulting condition,

$$\partial_\mu A^\mu = 0, \tag{5.37}$$

is called the *Lorentz condition*. A vector potential satisfying this condition is said to belong to the *Lorentz gauge*. This one condition effectively reduces the number of independent components of

A_μ from four to three. However, the Lorentz condition does not make A_μ unique; if A_μ satisfies the Lorentz condition, so will A'_μ as long as $\Box\Lambda(x) = 0$. Then by choosing $\Lambda(x)$ to satisfy

$$\frac{\partial \Lambda}{\partial t} = -\phi,$$

we have $\phi' = 0$, and thence, from (5.37), $\nabla\cdot\mathbf{A}' = 0$. Potentials satisfying this additional condition,

$$\phi = 0, \qquad \nabla\cdot\mathbf{A} = 0, \tag{5.38}$$

are said to belong to the *radiation* or *Coulomb gauge*. In this gauge there are clearly only two independent components of A_μ as happens in the real world.

Radiation gauge quantization

Since in the radiation gauge $\phi = 0$, the Maxwell's equations becomes

$$\Box\mathbf{A} = 0.$$

This is the Klein-Gordon equation for a massless field, and we write its solution as

$$\mathbf{A}(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \sum_{\lambda=1}^{2} \varepsilon^{(\lambda)}(k) \left[a^{(\lambda)}(k)e^{-ikx} + a^{(\lambda)\dagger}(k)e^{ikx}\right] \qquad (k^2 = 0,\ \ \omega_k = |\mathbf{k}|)$$

where $\varepsilon^{(\lambda)}$ are called *polarisation vectors* which verify

$$\mathbf{k}\cdot\,\varepsilon^{(\lambda)}(k) = 0.$$

In analogy with the quantization of KG field, we define the conjugate momentum fields

$$\pi^i = \frac{\partial \mathcal{L}}{\partial \dot{A}_i} = E^i, \tag{5.39}$$

and we impose the equal time commutation relations

$$\left[A^i(x), A^j(x')\right] = \left[E^i(x), E^j(x')\right] = 0 \tag{5.40}$$

$$\left[A^i(x), E^j(x')\right] = i\int \frac{d^3k}{(2\pi)^3}\left(\delta^{ij} - \frac{k^i k^j}{\mathbf{k}^2}\right) e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}. \tag{5.41}$$

The commutation relations for the operators $a^{(\lambda)}(k)$ and $a^{(\lambda)\dagger}(k)$ are

$$[a^{(\lambda)}(k), a^{(\lambda')\dagger}(k')] = 2\omega_k(2\pi)^3\delta_{\lambda\lambda'}\delta^3(\mathbf{k}-\mathbf{k}'),$$

$$[a^{(\lambda)}(k), a^{(\lambda')}(k')] = [a^{(\lambda)\dagger}(k), a^{(\lambda')\dagger}(k')] = 0$$

These commutation relations have the same form as those for the scalar field, and have the same interpretation as annihilation and creation operators for *photons*. The field energy (in the radiation gauge) is given by

$$\begin{aligned} H &= \frac{1}{2}\int d^3x \left(\mathbf{E}^2 + \mathbf{B}^2\right) = \int d^3x \left(\dot{\mathbf{A}}^2 + (\nabla\times\mathbf{A})^2\right) \\ &= \sum_\lambda \int \frac{d^3k}{4(2\pi)^3}[a^{(\lambda)}(k)a^{(\lambda)\dagger}(k) + a^{(\lambda)\dagger}(k)a^{(\lambda)}(k)], \end{aligned} \tag{5.42}$$

and on normal-ordering to remove the vacuum energy, we have

$$H = \sum_{\lambda} \int \frac{d^3k}{2(2\pi)^3} [a^{(\lambda)\dagger}(k) a^{(\lambda)}(k)].$$

Lorentz gauge quantization

Since our aim is to retain covariance, all four components of A_μ and of π_ν, will obey the covariant commutation relations

$$[A_\mu(\mathbf{x}, t), \pi_\nu(\mathbf{x}', t)] = i g_{\mu\nu} \delta^3(\mathbf{x} - \mathbf{x}'), \tag{5.43}$$

$$[A_\mu(\mathbf{x}, t), A_\nu(\mathbf{x}', t) = [\pi_\mu(\mathbf{x}, t), \pi_\nu(\mathbf{x}', t)] = 0, \tag{5.44}$$

where $g_{\mu\nu}$ is the Minkowski metric tensor, and

$$\pi^\mu = \frac{\partial \mathcal{L}}{\partial \dot{A}_\mu} \tag{5.45}$$

We immediately meet a problem because, with the Lagrangian density (5.34)

$$\pi^0 = \frac{\partial L}{\partial \dot{A}_0} = 0, \tag{5.46}$$

so it is not possible to satisfy (5.43) for A_0. Therefore, we need to change the Lagrangian, but the new Lagrangian should give the Maxwell's equations. The Euler-Lagrange equations with the lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2$$

gives

$$\Box A_\mu = 0 \tag{5.47}$$

as desired. But when we calculate π^0

$$\pi^0 = \frac{\partial \mathcal{L}}{\partial \dot{A}_0} = -\partial_\mu A^\mu,$$

which vanishes in the Lorentz gauge! The way out of this dilemma is to postulate that the Lorentz condition does not hold as an operator identity. Instead, we impose the weaker requirement that, for physical states $|\psi\rangle$, $\partial_\mu A^\mu$ has vanishing expectation value

$$\langle \psi | \partial_\mu A^\mu | \psi \rangle = 0 \tag{5.48}$$

The solution of (5.47) is clearly

$$A_\mu(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \sum_{\lambda=0}^{3} \varepsilon_\mu^{(\lambda)}(k) \left[a^{(\lambda)}(k) e^{-ikx} + a^{(\lambda)\dagger}(k) e^{ikx} \right].$$

Here the four polarisation 4-vectors $\varepsilon^{(\lambda)}$ have a Lorentz-invariant normalization

$$\varepsilon_\mu^{(\lambda)} \varepsilon^{(\lambda')} = \eta^{\lambda\lambda'}. \tag{5.49}$$

It is even too severe a condition to demand that physical state $|\psi\rangle$ should satisfy $\partial_\mu A^\mu |\psi\rangle = 0$ because the creation operators. We adopt the requeriement only for the part that contains annihilation operators and we have

$$\sum_{\lambda=0}^{\lambda=3} k^\mu \varepsilon_\mu^{(\lambda)} a^{(\lambda)}(k) |\psi\rangle = 0 \tag{5.50}$$

That is, the electromagnetic field only has two degrees of freedom.

The new (and covariant) commutation relations for the operators a and $a^\dagger$ are

$$\left[a^{(\lambda)}(k), a^{(\lambda')\dagger}(k')\right] = -\eta^{\lambda\lambda'} 2k_0 (2\pi)^3 \delta^3(\mathbf{k} - \mathbf{k}') \tag{5.51}$$

Chapter 6

Path integral formulation of quantum field theory

6.1 Path integrals in quantum mechanics

Consider a quantum mechanical system with one degree of freedom. The eigenstates of the position operator are introduced as follows

$$X_H(t)|x\rangle = x|x\rangle \qquad \text{Heisenberg picture}$$

$$X_S|x,t\rangle = x|x,t\rangle \qquad \text{Schrödinger picture}$$

with the relation

$$|x\rangle = \exp\left[-(i/\hbar)Ht\right]|x,t\rangle$$

where H denotes the Hamiltonian of the system. The matrix element

$$\langle x',t'|x,t\rangle = \langle x'|\exp\left[-(i/\hbar)H(t'-t)\right]|x\rangle \tag{6.1}$$

corresponds to the transition from the eigenstate $|x\rangle$ at the time t to the state $|x'\rangle$ at the time t', and is a Green's function.

The matrix element (6.1) we shall first represent as a multipli integral which shall then be used to define the funtional integral by a limitting procedure. First we divide the time interval $(t'-t)$ into $(n+1)$ equal parts of lenght ε

$$\begin{aligned} t' &= (n+1)\varepsilon + t \\ t_j &= j\varepsilon + t \qquad (j = 1,\ldots,n) \end{aligned} \tag{6.2}$$

Next, we use the completeness relation at each times t_j:

$$\int dx_j |x_j,t_j\rangle\langle x_j,t_j| = 1 \tag{6.3}$$

together with

$$\langle x_j,t_j|x_{j-1},t_{j-1}\rangle = \left\langle x_j \left| \exp\left(-\frac{i}{\hbar}\varepsilon H\right) \right| x_{j-1}\right\rangle = \langle x_j|x_{j-1}\rangle - \frac{i\varepsilon}{\hbar}\langle x_j|H|x_{j-1}\rangle + O(\varepsilon^2) \tag{6.4}$$

where $x_0, x_{n+1}, t_0, t_{n+1}$ are to be understood as x, x', t, t' respectively. Choosing the Hamiltonian $H = H(P, X)$ to be of the form $H = f(P) + g(X)$ we can write

$$\langle x_j|H|x_{j-1}\rangle = \int dp_j \langle x_j|p_j\rangle\langle p_j|H|x_{j-1}\rangle = \int \frac{dp_j}{2\pi\hbar} \exp\left[\frac{i}{\hbar}p_j(x_j - x_{j-1})\right] H(p_j, x_{j-1}) \tag{6.5}$$

where $H(p, x)$ is now the classical c-number Hamiltonian. Using (6.5), (6.4) becomes

$$\begin{aligned}\langle x_j, t_j|x_{j-1}, t_{j-1}\rangle &= \int \frac{dp_j}{2\pi\hbar} \exp\left[\frac{i}{\hbar}p_j(x_j - x_{j-1})\right]\left[1 - \frac{i}{\hbar}\varepsilon H(p_j, x_{j-1})\right] + O(\varepsilon^2) \\ &= \int \frac{dp_j}{2\pi\hbar} \exp\left[\frac{i}{\hbar}p_j(x_j - x_{j-1}) - \frac{i}{\hbar}\varepsilon H(p_j, x_{j-1})\right] + O(\varepsilon^2)\end{aligned} \tag{6.6}$$

and we obtain the following expression for the matrix element (6.1)

$$\langle x', t'|x, t\rangle = \lim_{n\to\infty} \int \prod_{j=1}^{n} dx_j \int \prod_{j=1}^{n+1} \frac{dp_j}{2\pi\hbar} \exp\left\{\frac{i}{\hbar}\sum_{j=1}^{n+1}[p_j(x_j - x_{j-1}) - H(p_j, x_{j-1})(t_j - t_{j-1})]\right\} \tag{6.1}$$

where the limit $n \to \infty \quad (\varepsilon \to 0)$ has been taken and the $O(\varepsilon^2)$ terms neglected. This result we shall write in the compact form

$$\langle x', t'|x, t\rangle = \int \frac{\mathcal{D}x\mathcal{D}p}{2\pi\hbar} \exp\left\{\frac{i}{\hbar}\int_t^{t'} [p\dot{x} - H(p, x)]d\tau\right\} \tag{6.7}$$

$\int(\mathcal{D}x\mathcal{D}p/2\pi\hbar) \equiv \int \prod_\tau (dx(\tau)dp(\tau))/(2\pi\hbar)$ is called a functional integration over all phase space, with the boundary conditions $x(t) = x$, $x(t') = x'$ implied in this case. Equation (6.7) is the path integral representation of $\langle x', t'|x, t\rangle$.

If the Hamiltonian is of the simple form

$$H = \frac{P^2}{2m} + V(X) \tag{6.8}$$

it is convenient to perform the momentum integrations and the final result has the form of a functional integral over configuration space

$$\langle x', t'|x, t\rangle = \frac{1}{N}\int \frac{\mathcal{D}x}{2\pi\hbar} \exp\left\{\frac{i}{\hbar}S[x]\right\}. \tag{6.9}$$

Here $S[x] = \int_t^{t'} L(x, \dot{x})d\tau$ is the action integral over the trajectory $x(\tau)$ where $L(x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x)$ is the Lagrangian and the normalization factor N is given by

$$\frac{1}{N} = \int \mathcal{D}p \exp\left(-\frac{i}{\hbar}\int_t^{t'} \frac{p^2}{2m}d\tau\right) \tag{6.10}$$

Starting with the canonically quantized theory described by the Hamiltonian (6.8) we have derived path integral representation (6.9). We can use another approach, namely, to define the quantum theory by the functional integral (6.9) i.e. we can choose the path integral formulation as the quantization prescription for a system with the classical Hamiltonian in the form (6.8).

6.2 Path integral formulation of quantum field theory

In field theory, the trajectory $x(t)$ is replaced by a field function $\Phi(\mathbf{x}, t)$. The degrees of freedom are now labelled by the continuous index $\mathbf{x}$; the number of degrees of freedom is obviously infinite. To define the appropiate path integral one can start from a multiple integral on a discrete and, for a beginning, finite lattice of space-time points. This amounts to defining the quantum field theory as a limit of a theory with only a finite number of degrees of freedom.

The limit of an infinite lattice, related to the thermodynamical limit of statistical mechanics, already defines a theory with infinite number of degrees of freedom. However, this lattice theory has not enough space-time invariance and a continuous theory must be defined. The latter limit is accompanied by infinities, the 'UV divergences' of quantum field theory. The definition of the functional integral in quantum field theory is thus more ambiguous than in the case of quantum mechanics. Nevertheless, the functional formalism in quantum field theory is of great heuristic value. It is very convenient tool for studying perturbation theory and allows a natural description of some non-perturbative phenomena.

The quantum field theory is usually formulated in terms of the vacuum expectation values of the chronologically ordered products of the field operators, the Green's functions

$$G^{(n)}(x_1, \ldots, x_n) = \langle 0|T\Phi(x_1) \ldots \Phi(x_n)|0\rangle \tag{6.11}$$

We postulate the following path integral representation

$$G^{(n)}(x_1, \ldots, x_n) \sim \int \mathcal{D}\Phi \, \Phi(x_1) \ldots \Phi(x_n) \exp\left[(i/\hbar) \int d^4x \, \mathcal{L}\right] \tag{6.12}$$

$\mathcal{D}\Phi$ denotes integration over all functions $\Phi(\mathbf{x}, t)$ of space and time, because, for each value of $\mathbf{x}$, $\Phi(\mathbf{x}, t)$ corresponds to a separate degree of freedom; $\mathcal{L}$ is the Lagrangian density.

If the time axis is rotated, $(t_i = -i\tau_i)$, the result is an *Euclidean Green's function*. The latter has a particularity convenient in the path integral representation, because the weight factor in the integrand: $\exp(-S_E/\hbar)$ is then non-negative. This Euclidean path integral formalism can be used to define the Minkowski space Green's functions by an analytic continuation of the Euclidean ones.

It is convenient to normalize the Green's functions by factorizing out the vacuum amplitude

$$\begin{aligned} G^{(n)}(x_1, \ldots, x_n) &= \langle 0|T\Phi(x_1) \ldots \Phi(x_n)|0\rangle \,/\, \langle 0|0\rangle \\ &= N \int \mathcal{D}\Phi \, \Phi(x_1) \ldots \Phi(x_n) \exp\left[(i/\hbar) \int d^4x \, \mathcal{L}\right] \end{aligned} \tag{6.13}$$

The Green's functions are given by the functional derivatives of the functional $W[J]$ equivalent to the vacuum transition amplitude in presence of the external source $J(x)$

$$W[J] = N \int \mathcal{D}\Phi \exp\left\{(i/\hbar) \int d^4x \, [\mathcal{L} + \hbar J(x)\Phi(x)]\right\} \tag{6.14}$$

Expanding in powers of J we can rewrite $W[J]$ as follows

$$W[J] = \sum_{n=0}^{\infty} \frac{i^n}{n!} \int dx_1 \ldots dx_n \, G^n(x_1, \ldots, x_n) J(x_1) \ldots J(x_n) \tag{6.15}$$

Consequently

$$G^n(x_1 \dots x_n) = \left(\frac{i^n}{n!}\right)^n \left.\frac{\delta^n}{\delta J(x_1)\dots\delta J(x_n)}\right|_{J\equiv 0} W[J] \tag{6.16}$$

The Green's functions can also be considered as the analytic continuation of those obtained from the generating functional defined in the Euclidean space with $x_0 = -i\tilde{x}_0$ and $\tilde{x}_0$ is real.

6.3 Introduction to perturbation theory

We shall discuss first de simple case of a scalar field theory described by the Lagrangian density

$$\mathcal{L} = \frac{1}{2}\left[\partial_\mu\Phi(x)\partial^\mu\Phi(x) - m^2\Phi^2\right] - \frac{\lambda}{4!}\Phi^4(x) \tag{6.17}$$

where x denotes four coordinates in Minkowski space. We seek a method for calculating an arbitrary Green's function in such theory. For that porpose it is convenient to use the functional formulation of the theory.

We recall at this point that, so far, we have been able to calculate exactly the generating functional $W_0[J]$ for a theory of non-interacting scalar fields with the action S_0 given by

$$S_0 = \int d^4x\, \mathcal{L}_{free} = \frac{1}{2}\int d^4x\, \left(\partial_\mu\Phi\partial^\mu\Phi - m^2\Phi^2\right) \tag{6.18}$$

But the exact solution is not known for the full theory and the method to be used for calculating the Green's functions is a perturbative expansion in terms of powers of S_1 defined as follows

$$S_I = S - S_0 \tag{6.19}$$

where

$$S = \int d^4x\, \mathcal{L} \tag{6.20}$$

Such an expansion expansion should be understood as an expansion in a neighbourhood of the vacuum state for which $\Phi(x) = 0$ so that the action S_I can be regarded as a small parameter. $W[J]$ can be rewritten as

$$W[J] = \frac{\exp\left\{\frac{i}{\hbar}S_I\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right\}W_0[J]}{\left.\exp\left\{\frac{i}{\hbar}S_I\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right\}W_0[J]\right|_{J\equiv 0}} = N\exp\left\{\frac{i}{\hbar}S_I\left[\frac{1}{i}\frac{\delta}{\delta J}\right]\right\}W_0[J] \tag{6.21}$$

The perturbation series is generated by spanding the exponential factor $\exp\left\{(i/\hbar)\, S_I\left[(1/i)\,(\delta/\delta J)\right]\right\}$ in powers of S_I and performing the functional differentations as indicated. In perturbations theory one gets, therefore, the following general formula for the Green's functions

$$G^{(n)}(x_1,\dots,x_n) = \frac{\int D\Phi\, \Phi(x_1)\dots\Phi(x_n)\left[\sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{i}{\hbar}S_I\right)^N\right]\exp\left(\frac{i}{\hbar}S_0\right)}{\int D\Phi\left[\sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{i}{\hbar}S_I\right)^N\right]\exp\left(\frac{i}{\hbar}S_0\right)} \tag{6.22}$$

Green's functions for free scalar field

The generating fuctional $W_0[J]$ for the free scalar field is given by

$$W_0[J] = N \int \mathcal{D}\Phi \exp\left\{(i/\hbar)\int d^4x\,[\mathcal{L} + \hbar J(x)\Phi(x)]\right\} \tag{6.23}$$

Using the identity (the surface term vanishes if $\phi \to 0$ at infinity)

$$\int \partial_\mu \Phi \partial^\mu \Phi d^4x = \int \partial_\mu \left(\Phi \partial^\mu \Phi\right) d^4x - \int \Phi \Box \Phi d^4x,$$

and the Green's function $G(x-y)$ for the free classical Klein-Gordon equation (the $i\varepsilon$ term dictates the path of integration round the poles at $k_0 = \pm\sqrt{k^2 - m^2}$)

$$G(x-y) = -\frac{1}{2\pi^4}\int d^4k \frac{\exp[-ik(x-y)]}{k^2 - m^2 + i\varepsilon} \qquad \varepsilon \to 0, \tag{6.24}$$

we have

$$W_0[J] = N \exp\left[\frac{1}{2} i\hbar \int d^4x d^4y J(x) G(x-y) J(y)\right] \tag{6.25}$$

Φ^4 theory

The Lagrangian density with ϕ^4 interaction term is given by

$$\mathcal{L} = \frac{1}{2}\partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} m^2 \Phi^2 - \frac{\lambda}{4!}\Phi^4 = \mathcal{L}_0 + \mathcal{L}_{\text{int}} \tag{6.26}$$

where

$$\mathcal{L}_{\text{int}} = \frac{\lambda}{4!}\Phi^4 \tag{6.27}$$

The generating functional with interaction is given by

$$W[J] = N \exp\left\{\frac{i\lambda}{\hbar\, 4!}\int \left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4 d^4x\right\} \exp\left[\frac{1}{2} i\hbar \int d^4x d^4y J(x) G(x-y) J(y)\right] \tag{6.28}$$

Expanding in powers of λ and using the so-called Feynman propagator

$$\Delta_F(x-y) = -G(x-y)/\hbar, \tag{6.29}$$

we have to order λ

$$\left(\frac{1}{i}\frac{\delta}{\delta J(z)}\right)^4 \exp\left[-\frac{1}{2} i \int d^4x d^4y J(x) \Delta_F(x-y) J(y)\right] =$$

$$\hbar^4 \left\{-3[\Delta_F(0)]^2 + 6i\Delta_F(0)\left[\int d^4x \Delta_F(z-x) J(x)\right]^2 + \left[\int d^4x \Delta_F(z-x) J(x)\right]^4\right\} \cdot \tag{6.30}$$

$$\exp\left[-\frac{1}{2} i \int d^4x d^4y J(x) \Delta_F(x-y) J(y)\right]$$

6.3.1 Feynman's Rules

We can extract graphical rules from Feynman rules, by which we can almost by inspection construct Green's functions of arbitrary complexity. With an interaction Lagrangian density given by $\lambda\phi^4/4!$, the matrix element $\mathcal{M}_{fi}$ (necessary to compute differential cross sections) can be calculated as follows:

1. Draw all possible connected, topologically distinct driagrams, including loops, with n external legs. Ignore vacuum-to-vacuum graphs.
2. For each internal line, associate a propagator given by:

$$i\Delta_F(p) = \frac{i}{p^2 - m^2 + i\epsilon}$$

3. For each vertex, associate the factor $-i\lambda$.
4. For each internal momentum corresponding to an internal loop, associate an integration factor:
$$\int \frac{d^4p}{(2\pi)^4} \tag{6.31}$$
5. Divide each graph by an overall symmetry factor S corresponding to the number of ways one can permute the internal lines and vertices, leaving the external lines fixed.
6. Momentum is conserved at each vertex.

Of course, the Feynman's rules are a bit more complicated for gauge theories.

6.4 Gauge and ghost fields

Consider a path integral over gauge field A_μ, corresponding to a physical, that is, gauge-invariant quantity

$$\int DA_\mu \, f\left(A_\mu\right) \exp\left(i \int d^4x \mathcal{L}\right) \tag{6.32}$$

For brevity we write A_μ instead of A^α_μ, where α is the gauge group index. $\mathcal{L}$ is the Lagrangian density and $f(A_\mu)$ denotes a gauge-invariant functional depending on the physical quantity under consideration. The integration measure DA_μ we assume to be invariant under gauge transformations. That is, it must have the following property

$$DA_\mu = DA^g_\mu \tag{6.33}$$

where g is an arbitrary transformation from the gauge group. A^g_μ denotes the result of this transformation when applied to A_μ.

The path integral in (6.32) runs over all possible configuration A_μ, which implies multiple counting of the physically equivalent configurations (equivalent up to a gauge transformation). Let us divide the configuration space $\{A_\mu(x)\}$ into the equivalence classes $\{A^g_\mu(x)\}$ of the gauge group. An orbit of the group includes all the field configurations which result when all possible transformations g from the gauge group G are applied to a given initial configuration $A_\mu(x)$. The integrand of (6.32) is constant along any orbit of the gauge group. Consequently, the

integral as it stands is proporcional to an infinite constant (the volume of the total gauge group). This is not an important difficulty in itself, because the infinity can always be cancelled by a normalization constant. The problem appears when we want to calculate (6.32) perturbatively because local gauge symmetry implies that the quadratic part of the gauge field action density has zero eigenvalues and therefore cannot be inverted in the configuration space $\{A_\mu(x)\}$ so that the propagator of the gauge field cannot be defined. One way to resolve this problem is to apply perturbation theory to the functional integral over the coset space of the orbits of the group (i.e, over the physically distinct field configurations) which follows from (6.32) after the infinite constant has been factorized out. We shall now describe a procedure by means of which this factorization can be achieved.

Let $\mathcal{D}$ denote an invariant measure on the gauge group G

$$\mathcal{D}g = \mathcal{D}(gg'); \qquad \mathcal{D}g = \prod_x dg(x) \tag{6.34}$$

and let us introduce a functional $\Delta[A_\mu]$ defined by the following equation

$$1 = \Delta[A_\mu] \int \mathcal{D}g \delta[F[A^g_\mu]]. \tag{6.35}$$

here $\delta[f(x)]$ represents the product of the usual Dirac δ-functions: $\prod_x \delta(f(x))$, one each space-time point. As for the functional $F[A_\mu]$, we assume that the equation

$$F[A^g_\mu] = 0 \tag{6.36}$$

has exactly one solution, g_0, for any initial field A_μ. We obtain, after formal manipulations

$$\Delta^{-1}[A_\mu] = \int \mathcal{D}F \left(\det \frac{\delta F[A^g_\mu]}{\delta g} \right)^{-1} \delta[F] \tag{6.37}$$

That is

$$\Delta[A_\mu] = \det \left. \frac{\delta F[A^g_\mu]}{\delta g} \right|_{F[A^g_\mu]=0} \tag{6.38}$$

is usually called the *Faddeev-Popov determinant*. $\Delta^{-1}[A_\mu]$ is invariant under gauge transformations.

Our aim is to replace integration over all field configurations by integration restricted to the hypersurface $F[A_\mu] = 0$. In this case each orbit will contribute only one field configuration and we shall have an integration over physically distinct fields. Inserting (6.35) under the path integral in (6.32) and change the order of integrations. Then we have

$$\int \mathcal{D}g \int \mathcal{D}A_\mu \Delta[A_\mu] f(A_\mu) \delta[F[A^g_\mu]] \exp\{iS[A_\mu]\} \tag{6.39}$$

An important observation is that the complete expression under $\int \mathcal{D}g$ integral is independent of g. The group integration $\int \mathcal{D}g$ factories out to produce an infinite constant: the volumen of the full gauge group. Now, in the $F[A_\mu] - C(x) = 0$ gauge (6.39) becomes

$$\left(\int \mathcal{D}g \right) \int \mathcal{D}A_\mu \Delta[A_\mu] \delta[F[A^g_\mu] - C(x)] f(A_\mu) \exp\{iS[A_\mu]\}. \tag{6.40}$$

Being gauge invariant, this is obviously independent of $C(x)$. We can integrate (6.40) functionally over $\int \mathcal{D}C$ with an arbitrary weight functional $G[C]$; the result will differ from (6.40) only

by overall normalization constant. Observing that δ-functional in (6.40) is the only C-dependent term, we obtain

$$N \int \mathcal{D}A_\mu \Delta[A_\mu] f(A_\mu) \exp\{iS[A_\mu]\} G[F[A_\mu]] \tag{6.41}$$

A choice for $G[C]$ is

$$G[C] = \exp\left\{-\frac{i}{2\alpha}\int d^4x [C(x)]^2\right\} \tag{6.42}$$

where α is a real constant.

It is convenient to use the gauge invariance property of $\Delta[A_\mu]$ to choose A_μ which satisfies the gauge condition $F[A_\mu] = 0$. Then in (6.38) we can replace the constraint $F[A^g_\mu] = 0$ by $g = 1$, which simplifies the practical calculations

$$\Delta\left[A_\mu\right] = \det\left.\frac{\delta F[A^g_\mu]}{\delta g}\right|_{g=1}; \qquad F[A_\mu] = 0 \tag{6.43}$$

Near $g = 1$ we only have to deal with the infinitessimal transformations: $U(\Theta) = 1 - iT^\alpha \Theta^\alpha(x)$ (where $\Theta^\alpha(x) \ll 1$) and the invariant group measure $\mathcal{D}g$ takes the simple form $\Delta\Theta \equiv \prod_{\alpha,x} d\Theta^\alpha(x)$.

We can now rewrite (6.43) in a more explicit form, with all relevant indices

$$\Delta\left[A_\mu\right] = \det\left.\frac{\delta F\left(x,[A^g_\mu]\right)}{\delta\Theta^\beta(y)}\right|_{\Theta=0}; \qquad F^\alpha\left(x,[A_\mu]\right) = 0 \tag{6.44}$$

A^Θ_μ stands for A^g_μ and α, β are the gauge group indices. We have to calculate a determinant of a matrix in both space-time and the group indices $M^{\alpha\beta}(x,y)$. This matrix appears in the expansion of $F^\alpha\left(x,[A^\Theta_\mu]\right)$ in powers of the infinitessimal parameters $\Theta^\beta(y)$

$$F^\alpha\left(x,[A^\Theta_\mu]\right) = F^\alpha\left(x,[A_\mu]\right) + \int d^4y M^{\alpha\beta}(x,y)\Theta^\beta(y) + \cdots \tag{6.45}$$

so that

$$M^{\alpha\beta}(x,y) = \left.\frac{\delta F\left(x,[A^g_\mu]\right)}{\delta\Theta^\beta(y)}\right|_{\Theta=0} \tag{6.46}$$

and

$$\Delta[A_\mu] = \det M; \qquad F[A_\mu] = 0\ . \tag{6.47}$$

The standard method of dealing with the Faddeev-Popov determinant $\det M^{\alpha\beta}(x,y)$ is to replace it by an additional functional integration over some auxiliary complex fields $\eta(x)$ (ghost fields) which are Grassmann variables

$$\det M^{\alpha\beta}(x,y) = C\int \mathcal{D}\eta\mathcal{D}\bar\eta \exp\left[i\int d^4x d^4y\, \bar\eta_\alpha(x) M^{\alpha\beta}(x,y)\eta_\beta(y)\right] \tag{6.48}$$

where C is some constant that includes the factor $(-1/g)$. Now putting $\Delta[A_\mu] = \det M$ we get

$$W = N\int \mathcal{D}A_\mu\mathcal{D}\eta\mathcal{D}\bar\eta \exp\left[i\int d^4x\left(\mathcal{L} - \frac{1}{2\alpha}F^2 - \bar\eta_\alpha(x)M^{\alpha\beta}(x)\eta_\beta(x)\right)\right] \tag{6.49}$$

This is equivalent to replacing the original Lagrangian density by

$$\mathcal{L}_{\text{eff}} = \mathcal{L} - \frac{1}{2\alpha}(F[A])^2 - \bar\eta_\alpha M^{\alpha\beta}\eta_\beta = \mathcal{L} + \mathcal{L}_{GF} + \mathcal{L}_{FPG} \tag{6.50}$$

$\mathcal{L}_{GF}$ is the gauge-fixing term, and $\mathcal{L}_{FPG}$ is the Faddeev-Popov ghost term. The Grassmann fields η and $\bar\eta$ are called *ghost fields*.

6.5 Renormalization

One of the serious complications found in quantum field theory (QFT) is the fact that the theory is naively divergent. We can see that the divergences found in QFT were, in some sense, inevitable. In the transition from quantum mechanics to QFT, we made the transition from a finite number of degrees of freedom to an infinite number. Because of this, we must continually sum over an infinite number of internal modes in loop integrations, leading to divergences. In the approach to field theory based on perturbation theory it is imperative to make sense of the perturbation series. In order for a field theory to be at all believable, the problems raised by divergences must be satisfactorily resolved.

One approach is to proceed order by order in perturbation theory and show that at each order the quantities of physical interest can be renormalized to finite values. If this is possible to all orders then theory is *renormalizable*.

Since that time, there have been two important developments in renormalization theory. The first was the renormalization of *Quantum Electrodynamics* (QED) via the covariant formulation developed by Schwinger and Tomonaga and by Feynman (which were shown to be equivalent by Dyson). The second development was the proof by 't Hooft that spontaneously broken Yang-Mills theory was renormalizable, which led to the succesful application of QFT to the weak interactions.

The essential idea is that there is a set of "bare"physical parameters that are divergent, such as masses, coupling constants or Green's functions. However, these bare parameters are unmeasurable. The divergences of these parameters are chosen so that they cancel against the ultraviolet infinities coming from infinite classes of Feynman diagrams, which probe the small-distance behaviour of the theory. After these divergences have been absorbed by the bare parameters, we are left with the physical, renormalized parameters that are indeed measurable. Renormalization theory, then, is a set of rules or prescriptions where, after a finite number of redefinitions, we can render the theory finite to any order. The problem with renormalization lies in the details. We present only the basic components of renormalization theory, which occur in four essential steps:

1. *Power counting.* By simple counting the powers of p in any Feynman graph, we can, for large p, tell whether the integral diverges by calculating the degree of divergence of that graph: each boson propagator contributes p^{-2}, each fermion propagator contributes p^{-1}, each loop contributes a loop integration with p^4, and each vertes with n derivatives contributes at most n powers of p. If the overall power of p; that is, the degree of divergence D, is 0 or positive, then the graph diverges. By simple power counting arguments, we can then calculate rather quickly whether certain theories are non-renormalible, or whether they can be potentially renormalized.

2. *Regularization.* Manipulating divergent integrals is not well defined, so we need to cutoff the integration over d^4p. This formally renders each graph finite, order by order, and allows us to reshuffle the perturbation theory. At the end of the calculation, after we have rearranged the graphs to put all divergent terms into the physical parameters, we let the cutoff tend to infinity. We must also show that the resulting theory is independent of the regularization method.

3. *Counterterms or multiplicative renormalization.* Given a divergent theory that has been regularized, we can perform formal manipulations on the Feynman graphs to any order. Then there are at least two equivalent ways in which to renormalize the theory:

First, there is the method of *multiplicative renormalization*, pioneered by Dyson and Ward for QED, where we formally sum over an infinite series of Feynman graphs with a fixed number of external lines. The divergent sum is the absorbed into a redefinition of the coupling constants and masses in the theory. Since the bare masses and bare coupling constants are unmeasurable, we can assume they are divergent and that they cancel against the divergences of corresponding Feynman graphs, and hence the theory has absorbed all divergences at that level.

Second, there is the method of *counterterms*, pioneered by Bogoliubov, Parasiuk, Hepp, and Zimmerman (BPHZ), where we add new terms directly to the action to substract off the divergent graphs. The coefficients of these counterterms are chosen so that they precisely kill the divergent graphs. In a renormalizable theory, there are only a finite number of counterms needed to render the theory finite to any order. Furthermore, these counterterms needed to render the theory finite to any order. Furthermore, these counterterms are proportional to terms in the original action. Adding the original action with the counterterms give us a renormalization of the masses and coupling constants in the action. These are equivalent and give us simple criteria that are necessary (but not sufficient) to prove that a theory is renormalizable:

a. The degree of divergence D of any graph must be a function only of the number of external legs; that is, it must remain constant if we add more internal loops. This allows us to collect all N-point loop graphs into one term.
b. The number of classes of divergent N-point graphs must be finite. These divergences must cancel against the divergences contained whithin the bare parameters.

4. *Induction.* The last step in the proof of renormalizability is to use an induction argument. We assume the theory is renormalizable at the nth order in perturbation theory. Then we write down a recursion relation that allows us to generate the $n+1$st-order graphs in terms of the nth-order graphs. By proving the $n+1$st-order graphs are all finite, we can prove, using either multiplicative or counterterm renormalization, that the entire perturbation theory, order by order, is finite. All induction proofs ultimately rely on *Weinberg's theorem* which states that a Feynman graph converges if the degree of divergence of the graph and all its subgraphs is negative.

6.5.1 Renormalization group and renormalization group equation

We let R represent some (unspecified) renormalization scheme. If Γ_0 is an unrenormalized quantity and Γ_R is same quantity and Γ_R is same quantity renormalized by the scheme R, then

$$\Gamma_R = Z(R)\Gamma_0 \tag{6.51}$$

where $Z(R)$ represents some renormalization constant under the renormalization scheme R.

Let us now choose a different renormalization scheme R'. Since the unrenormalied quantity Γ_0 was independent of renormalization scheme, then

$$\Gamma_{R'} = Z(R')\Gamma_0 \tag{6.52}$$

Then the relation between these two renormalized quantities is given by

$$\Gamma_{R'} = Z(R', R)\Gamma_R \tag{6.53}$$

where

$$Z(R', R) \equiv Z(R'/Z(R). \tag{6.54}$$

Trivially, this satisfies a group multiplication law

$$Z(R'', R')Z(R', R) = Z(R'', R) \tag{6.55}$$

where the identity element is given by

$$Z(R, R) = 1. \tag{6.56}$$

In the technique of dimensional regularization it is necessary to introduce a new parameter μ. The renormalized 1PI (1-particle irreducible) function $\Gamma_R^{(n)}$ depends on μ, while the unrenormalized vertex functions are independent of μ, the renormalized ones are not. For example, in ϕ^4 theory, we have the following relationship between unrenormalized and renormalized quantities

$$\Gamma_0^{(n)}(p_i, g_0, m_0) = Z_\phi^{-n/2}\Gamma^{(n)}(p_i, g, m, \mu) \tag{6.57}$$

where μ is the substraction point, and we assume that have used some regularization scheme to render all expressions finite for the moment.

Now let us differentiate this via the dimensionless derivative $\mu(d/d\mu)$. We know that the unrenormalized bare quantity is independent of substraction point, so that the derivative acting on the unrenormalized quantity must, by construction, be zero:

$$0 = \mu\frac{\partial}{\partial\mu}\Gamma_0^{(n)} = \left(\mu\frac{\partial}{\partial\mu}Z_\phi^{-n/2}\right)\Gamma^{(n)} + Z_\phi^{-n/2}\left(\mu\frac{\partial}{\partial\mu}\Gamma^{(n)}\right) \tag{6.58}$$

We now use the chain rule. We choose as our independent variables μ, g, and m:

$$\frac{d}{d\mu} = \frac{\partial}{\partial\mu} + \frac{\partial g}{\partial\mu}\frac{\partial}{\partial g} + \frac{\partial m}{\partial\mu}\frac{\partial}{\partial m} \tag{6.59}$$

Let us make the following definitions (where we now take the limit as $\epsilon \to 0$):

$$\begin{aligned}
\beta(g) &\equiv \mu\frac{\partial g}{\partial\mu} \\
\gamma(g) &\equiv \mu\frac{\partial}{\partial\mu}\log\sqrt{Z_\phi} \\
m\gamma_m(g) &\equiv \mu\frac{\partial m}{\partial\mu}
\end{aligned} \tag{6.60}$$

With these definitions, we now have the compact expression:

$$\left(\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} - n\gamma(g) + m\gamma_m(g)\frac{\partial}{\partial m}\right)\Gamma^{(n)}(p_i, g, m, \mu) = 0 \tag{6.61}$$

This is the *renormalization group equation*, and it express how the renormalized vertex functions change when we make a change in the substraction point μ.

We can write down a similar equation expressing the invariance of Γ^n under a change of scale. Let $p_i \to tp_i$ Using dimensional arguments, the vertex function function behaves as

$$\Gamma^n(tp_i, g, m, \mu) = \mu^D F\left(g, \frac{t^2p_i^2}{m\mu}\right) \tag{6.62}$$

where D is the dimension of the vertex function. This implies that the vertex function obeys

$$\left(t\frac{\partial}{\partial t}\mu\frac{\partial}{\partial\mu} + m\frac{\partial}{\partial m} - D\right)\Gamma^{(n)} = 0 \tag{6.63}$$

Now let us eliminate the term $\mu(\partial\Gamma/\partial\mu)$ from thios equation using (6.61). Then we find

$$\left(-t\frac{\partial}{\partial t}+\beta(g)\frac{\partial}{\partial g}-n\gamma(g)+m(\gamma_m(g)-1)\frac{\partial}{\partial m}+D\right)\Gamma^{(n)}(tp_i,g,m,\mu)=0 \tag{6.64}$$

This equation expresses directly the effect on $\Gamma^{(n)}$ of scaling up momenta by a factor t. We can find a solution of (6.64) looking for solutions of the form

$$\Gamma^{(n)}(tp_i,m,g,\mu)=f(t)\Gamma^{(n)}(p_i,m,g,\mu). \tag{6.65}$$

Multiplying this by $t\dfrac{\partial}{\partial t}$ and using (6.65)

$$\left(-t\frac{\partial}{\partial t}+\frac{t}{f}\dot{f}+t\frac{m}{\partial t}\frac{\partial}{\partial m}+t\frac{g}{\partial t}\frac{\partial}{\partial g}\right)\Gamma^{(n)}(tp_i,g,m,\mu)=0. \tag{6.66}$$

Comparing (6.66) with (6.64) we have

$$\begin{aligned} t\frac{\partial g(t)}{\partial t}&=\beta(g(t)),\\ t\frac{\partial m}{\partial t}&=m(\gamma_m(g)-1),\\ \frac{t}{f}\dot{f}&=D-n\gamma(g). \end{aligned} \tag{6.67}$$

$g(t)$ is the *running constant coupling*. Integrating the last equation we have

$$f(t)=t^D\exp\left(-\int_1^t\frac{n\gamma(g(t))\,dt}{t}\right), \tag{6.68}$$

which, on substitution into (6.65) gives the solution to the equation (6.64) in terms of the running coupling constant $g(t)$ and running mass $m(t)$.

6.6 Asymptotic freedom of Yang-Mills theories

Now we examine some possible behaviours of $g(t)$ as $t\to\infty$, i.e, at large momentum, and assume that

$$t\frac{\partial g(t)}{\partial t}=\beta(g(t)) \tag{6.69}$$

is still valid there. Knowledge of the function $\beta(g)$ enables us to find $g(t)$. The zeros of β are called *fixed points*. If $\beta<0$, g decreases with increasing t, and when $t\to\infty$ $g=0$. This is known as *asymptotic freedom*.

It turns out that asymptotic freedom is a property possessed by all non-abelian gauge theories. We shall not study the general case, but confine ourselves, for definiteness and because of the physical relevance of QCD, to $SU(3)$ gauge symmetry.

If g is the Yang-Mills coupling constant (charge), whose physical and bare values are related by

$$g_B=g\mu^{\varepsilon/2}Z_1Z_2^{-1}Z_3^{-1/2} \tag{6.70}$$

Here Z_1 is the renormalisation constant for the quark-gluon-gluon vertex, Z_2 that for the quark wave function, and Z_3 that for the gluon wave function (self-energy).

$$Z_1 = 1 - \frac{g}{82\pi^2}\frac{13}{3}\ , \qquad Z_2 = 1 - \frac{g^2}{6\pi^2\varepsilon}\ , \qquad Z_3 = 1 + \frac{g^2}{8\pi^2\varepsilon}\left(5 - \frac{2n_F}{3}\right)$$

where n_F is the number of quark flavours of quark (probably six). In the limit $\varepsilon \to 0$

$$\beta(g) = \mu\frac{\partial g}{\partial \mu} = \frac{g^3}{16\pi^2}\left(-11 + \frac{2n_F}{3}\right) \tag{6.71}$$

As the number of quark flavours is $n_F \leq 16$, then $\beta < 0$ and g decreases with increasing masses (momentum) scale μ, so the theory is asymptotically free.

In summary, asymptotic freedom means that, roughly speaking, at shorter and shorter distances, the coupling constant decreases in size, so that the theory appears to be a free theory. Conversely, at larger and larger distances, the coupling constant increases, so that at a certain point perturbative calculations can no longer be trusted. Large coupling constants, in turn, imply that the quarks bind more tightly together, giving rise a confinement. This is called "infrared slavery", which is the counterpoint of asymptotic freedom.

Chapter 7

Application of Yang-Mills theory to the Standard Model

7.1 The Standard Model of the elementary particles

The *Stardard Model* (SM) constitutes one of the most succesful achievements in modern physics. The SM is a gauge theory, based on the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which describes strong, weak and electromagnetic interactions, via the exhange of the corresponding spin-1 gauge fields: eight massless gluons and one massless photon, respectively, for the strong and electromagnetic interactions, and three massive bosons, $W^{\pm}$ and Z, for the weak interaction. The fermionic matter content is given by six leptons (electron e^-, muon μ^-, tau τ^- and three neutrinos) and six quarks (up u, down d, charm c, strange s, bottom b and top t) and their corresponding antiparticles

$$\text{quarks:} \begin{pmatrix} u & d & c \\ s & b & t \end{pmatrix} \qquad \text{leptons:} \begin{pmatrix} e^- & \mu^- & \tau^- \\ \nu_e & \nu_\mu & \nu_\tau \end{pmatrix}, \tag{7.1}$$

which are organized in three families: (ν_e, e^-, u, d), (ν_μ, μ^-, c, s) and (ν_τ, τ^-, b, t). There is a basic classification of fundamental particles into those which experience the strong interaction, called *hadrons*, those which do not, called *leptons*, and, thirdly the quanta of the interaction fields. The hadrons are subdivided into *baryons*, which have half-odd integral spin, and *mesons*, with integral spin (in units of $\hbar$). According to the quark model, baryons are bound states of three quarks (qqq), and mesons are quark-antiquark ($q\bar{q}$) states.

The gauge symmetry is broken by the vacuum, which triggers the *Spontaneous Symmetry Breaking* (SSB) of the electroweak group to the electromagnetic subgroup. The SSB mechanism generates the masses of the weak gauge bosons, and give rise to the appearance of a physical scalar particle in the model, the so-called Higgs Boson. The fermion masses and mixings are also generated through the SSB.

The SM is a mathematically-consistent renormalizable theory which predicts or is consistent with all experimental facts. However, the theory has too much arbritrariness to be the final story. For example, the minimal version has 21 free parameters and most physicists believe that this is just too much for the fundamental theory. And we don't understand why only the electroweak part is chiral (parity-violating). Similarly, the SM incorporates, but does not explain, charge quantization (i.e., why all particles have charges which are multiples of $e/3$). We don't know why there are three fermion families, and why all matter under ordinary terrestrial

conditions can be constructed out of the fermions of the first family (ν_e, e^-, u, d). We don't know how to calculate theoretically the Higgs Boson mass (if it exists). And the SM does not includes gravity.

7.2 Quantum electrodynamics (QED)

Quantum electrodynamics (QED) is perhaps the best fundamental physical theory we have. The theory is formulated as a set of simple equations (Maxwell's equations and the Dirac equation) whose form is essentially determined by relativistic invariance. The quantum-mechanical solutions of these equations give detailed predictions of electromagnetic phenomena from macroscopic distances down to regions several hundred times smaller than the proton.

The QED Lagrangian density is given by

$$\mathcal{L}_{QED} = \mathcal{L}_{KIN} + \mathcal{L}_0 = -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x) + i\bar{\psi}(x)\gamma^\mu D_\mu\psi(x) - m\bar{\psi}(x)\psi(x). \tag{7.2}$$

where $\mathcal{L}_{KIN} = -\frac{1}{4}F_{\mu\nu}(x)F^{\mu\nu}(x)$ is the gauge-invariant kinetic term and $\mathcal{L}_0 = i\bar{\psi}(x)\gamma^\mu D_\mu\psi(x) - m\bar{\psi}(x)\psi(x)$ is the free Dirac fermion Lagrangian density with minimal coupling.

Here, $D_\mu \equiv \partial_\mu + ieA_\mu$ is the covariant derivative, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the usual electromagnetic field strength and e and m are parameters of the theory. The Lagrangian density $\mathcal{L}$ is invariant under local $U(1)$ transformations, i.e, is invariant under the gauge transformations

$$\begin{aligned} \psi'(x) &= \exp\{i\theta(x)\}\psi(x) \\ A'_\mu(x) &= A_\mu(x) - \frac{1}{e}\partial_\mu\theta(x) \end{aligned} \tag{7.3}$$

where $\theta(x)$ depends on space and time in a completely arbitrary way. A possible mass term for the gauge field, $\mathcal{L}_m = \frac{1}{2}m^2A_\mu A^\mu$, is forbidden because it would violate gauge invariance; therefore, the photon field is predicted to be massless.

The total Lagrangian gives rise to the well-known Maxwell equations:

$$\partial_\mu F^{\mu\nu} = J^\nu \equiv e\bar{\psi}\gamma^\nu\psi. \tag{7.4}$$

The energy-momentum tensor is given by

$$T^{\mu\nu} = F^{\mu\rho}F^\nu_\rho + \frac{1}{4}g^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta} \tag{7.5}$$

is conserved, symmetric, and gauge invariant.

7.2.1 Renormalization of QED

Over the years, a large number of renormalization programs have been developed, with various degrees of rigor, and each with their own advantages and disadvantages. Three proofs are

1. The original Dyson/Ward proof.
2. The BPHZ proof.
3. Proof based on the Callan-Symanzik equations.

7.3 Quantum chromodynamics (QCD)

The simple quark model was initially developed in early-1960s to account for the regularities observed in the hadron spectrum, with hadrons interpreted as bound states of localized but essentially non-interacting quarks.

Quarks of a given type (u, d, s, ...) possess an addicional label called R, G or B, which takes on three values.The quark Lagrangian is invariant under relabellings described by a group of tranformations

$$q \to Mq \ , \ q = \begin{pmatrix} q_R \\ q_G \\ q_B \end{pmatrix} \tag{7.6}$$

M es a 3×3 matrix, and may be either orthogonal or unitary (these matrices form groups - hermitian matrices do not). Which do we choose? We choose $SU(3)$ rather than $O(3)$ for two reasons: (a) If the colour group were $O(3)$, the diquark system could be a colour singlet, but diquarks are not found in nature; (b) $O(3)$ does not possess asymptotic freedom if the number of flavours exceeds two. M is then a unitary matrix, which, because we may substract out an overall phase, may be chosen to have unit determinant

$$q \to Uq \quad (\text{or} \quad q^\alpha_\beta \to (q^\alpha_f)' = U^\alpha_\beta q^\beta_f \,), \quad U^\dagger U = 1, \quad \det U = 1. \tag{7.7}$$

The quark field has two labels, the colour α and the flavour f. U may be written in the form

$$U = e^{iH}, \quad H = H^\dagger U, \quad \mathrm{Tr} H = 0$$

where H is Hermitian and the zero trace condition follows from $\det U = 1$. It may be seen that U has eight independent parameters ε_a, and therefore eight generators, which we denote $\lambda_a/2$, so we write

$$U = \exp\left(i\frac{\lambda_a}{2}\varepsilon_a\right) \tag{7.8}$$

where a summation over a from 1 to 8 is implied, and the eight matrix generators are the matrices of Gell-Mann (2.12). The matrices λ_a obey the commutation relations

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2}\right] = i f_{abc}\frac{\lambda_c}{2} \tag{7.9}$$

(with summation over c from 1 to 8). The quantities if_{abc} are the structure constants of the group and are totally antisymmetric in their indices. The only non-zero components are

$$\begin{aligned} &f_{123} = 1, \\ &f_{147} = -f_{156} = f_{246} = f_{257} = f_{345} = -f_{367} = \frac{1}{2}, \\ &f_{458} = f_{678} = \frac{\sqrt{3}}{2} \end{aligned} \tag{7.10}$$

According to the theory of chromodynamics, colour symmetry is a *gauge symmetry*, so the theory we have developed above may be taken over wholesale. In particular, there is a gauge potential, or rather eight gauge potentials which may be written in matrix form

$$A_\mu = A^a_\mu \frac{\lambda_a}{2} = \frac{1}{2}\begin{pmatrix} A^3_\mu + \frac{1}{\sqrt{3}}A^8_\mu & A^1_\mu - iA^2_\mu & A^4_\mu - iA^5_\mu \\ A^1_\mu + iA^2_\mu & -A^3_\mu + \frac{1}{\sqrt{3}}A^8_\mu & A^6_\mu - iA^7_\mu \\ A^4_\mu + iA^5_\mu & A^6_\mu + iA^7_\mu & -\frac{2}{\sqrt{3}}A^8_\mu \end{pmatrix}. \tag{7.11}$$

The gauge fields $G^a_{\mu\nu}$ may be written down

$$G^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f_{abc} A^b_\mu A^c_\nu \tag{7.12}$$

and

$$G_{\mu\nu}(x) = -\frac{i}{g}[D_\mu, D_\nu] = \frac{\lambda_a}{2} G^a_{\mu\nu}(x), \tag{7.13}$$

where

$$D_\mu \equiv \partial_\mu + igA_\mu. \tag{7.14}$$

They are gauge fields, and they are massless, since the presence of mass terms for gauge fields destroys the gauge invariance of the Lagrangian density. They go by the name of *gluons*. If the explanation of quark confinement is that colour is confined, and therefore that only colour singlet particles may appear in the free state, then gluons will also be confined and therefore unobservable. There should, however, be a combination of gluon fields invariant under $SU(3)$, and therefore in principle, observable. This is called glueball.

Under a gauge transformation

$$\begin{aligned} D_\mu &\to (D_\mu)' = U D_\mu U^\dagger, \\ A_\mu &\to (A_\mu)' = U A_\mu U^\dagger + \frac{i}{g} (\partial_\mu U) U^\dagger, \\ G_{\mu\nu} &\to (G_{\mu\nu})' = U G_{\mu\nu} U^\dagger, \end{aligned} \tag{7.15}$$

and the colour trace $\mathrm{Tr}(G^{\mu\nu} G_{\mu\nu})$ remains invariant.

The QCD Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} + \sum_f \left(i\bar{q}_f \gamma^\mu D_\mu q - m_{q_f} \bar{q}_f q_f \right) + \mathcal{L}_{GF} + \mathcal{L}_{FPG} \tag{7.16}$$

where $\mathcal{L}_{GF}$ and $\mathcal{L}_{FP}$ are the gauge-fixing term and the Faddeev-Popoc term (or ghost term)

$$\begin{aligned} \mathcal{L}_{GF} &= -\frac{1}{2} (F[A])^2 \\ \mathcal{L}_{FPG} &= -\bar{\eta}_\alpha(x) M^{\alpha\beta}(x) \eta_\beta(x) \end{aligned} \tag{7.17}$$

The most used gauges are the covariant gauge ($\partial_\mu A^{\mu a} = 0$) and the axial gauge ($n_\mu A^{\mu a} = 0$). The ghosts $\eta^\alpha(x)$ are scalar fields with Fermi statistics.

The field strength tensor $G_{\mu\nu}$ has a remarkable new property on account of the term $g f_{abc} A^b_\mu A^c_\mu$. Imposing the gauge symmetry has requiered that the kinetic energy term in $\mathcal{L}$ is not purely kinetic but includes an induced self-interaction between the gauge bosons.

7.4 The $\mathbf{SU(2)_L \otimes U(1)}$ theory of electroweak interactions

Though al hadrons and leptons experience the weak interaction, and hence, can undergo weak decays, they are often hidden by the much more rapid color a electromagnetic decays. The weak interactions is also responsible for the β-decay of atomic nuclei, which involves the transformation of a proton to a neutron (or vice-versa).

The simplest group to describe the weak interactions is $SU(2)$. Moreover, if we want to include the electromagnetic interactions we need and additional $U(1)$ group. The obvious symmetry group to consider is the $G = SU(2)_L \otimes U(1)$, where L refers to left-handed fields. The identification of the $U(1)$ group with electromagnetism does not work.

For simplicity, let us consider a single family of quarks, and introduce the notation

$$\psi_1(x) = \begin{pmatrix} u \\ d \end{pmatrix}_L , \qquad \psi_2(x) = u_R , \qquad \psi_3(x) = d_R \tag{7.18}$$

Our discussion will also be valid for the lepton sector, with the identification

$$\psi_1(x) = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L , \qquad \psi_2(x) = \nu_{eR} , \qquad \psi_3(x) = e^-_R \tag{7.19}$$

Let us consider the free Lagrangian density (Dirac Lagrangian with $m = 0$)

$$\mathcal{L}_0 = \sum_{j=1}^{3} i\bar{\psi}_j(x)\gamma^\mu \partial_\mu \psi_j(x) \tag{7.20}$$

$\mathcal{L}_0$ is invariant under global G transformations in flavour space:

$$\begin{aligned} \psi_1(x) &\to \psi_1'(x) = \exp\{iy_1\beta\}\, U_L \psi_1(x) \\ \psi_2(x) &\to \psi_2'(x) = \exp\{iy_2\beta\}\, U_L \psi_2(x) \\ \psi_3(x) &\to \psi_3'(x) = \exp\{iy_3\beta\}\, U_L \psi_3(x) \end{aligned} \tag{7.21}$$

where the $SU(2)_L$ transformation

$$U_L = \exp\left\{ i\frac{\sigma_i}{2}\alpha^i \right\} \qquad\qquad (i = 1, 2, 3) \tag{7.22}$$

only acts on the doublet field ψ_1 (σ_i are the Pauli matrices). The parameters y_i are called *hypercharges*, since the $U(1)_Y$ phase transformation is analogous to the QED one. The matrix transformation U_L is non-abelian as in QCD.

We can now require the Lagrangian to be also invariant under local $SU(2)_L \otimes U(1)_Y$ gauge transformations, i.e., with $\alpha^i = \alpha^i(x)$ and $\beta = \beta(x)$. In order to satisfay this symmetry requirement, we need to change the fermion derivatives by covariant objects. Since we have now four gauge parameters, $\alpha^i(x)$ and $\beta(x)$, four different gauge bosons are needed:

$$\begin{aligned} D_\mu \psi_1(x) &= \left[\partial_\mu + ig\widetilde{W}_\mu(x) + ig' y_1 B_\mu(x) \right] \psi_1(x), \\ D_\mu \psi_2(x) &= \left[\partial_\mu + ig' y_2 B_\mu(x) \right] \psi_2(x), \\ D_\mu \psi_3(x) &= \left[\partial_\mu + ig' y_3 B_\mu(x) \right] \psi_3(x), \end{aligned} \tag{7.23}$$

where

$$\widetilde{W}_\mu(x) = \frac{\sigma^i}{2} W^i_\mu(x) \tag{7.24}$$

denotes a $SU(2)_L$ matrix field. Thus we have the correct number of gauge fields to describe the $W^\pm$, Z and γ.

The gauge transformations for the fields are:

$$\begin{aligned} B'_\mu(x) &= B_\mu - \frac{1}{g}\partial_\mu\beta(x) \\ \widetilde{W'}_\mu &= U_L\widetilde{W}_\mu U_L^\dagger(x) + \frac{i}{g}\partial_\mu U_L(x)U_L^\dagger(x), \end{aligned} \tag{7.25}$$

where $U_L = \exp\left\{i\frac{\sigma_i}{2}\alpha^i(x)\right\}$. The transformation of B_μ is identical to the one obtained in QED for the photon, while the $SU(2)_L$ W^i_μ fields transform in a way analogous to the gluon fields of QCD. Note that the ψ_j couplings to B_μ are completely free as in QED, i.e., the hypercharges y_j can be arbitrary parameters. Since the $SU(2)_L$ commutation relation is non-linear, this freedom does not exist for the W'_μ: there is only a unique $SU(2)_L$ coupling g.

Now, we introduce the corresponding fields strengths:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{7.26}$$

$$\widetilde{W}_{\mu\nu} = -\frac{i}{g}\left[\left(\partial_\mu + ig\widetilde{W}_\mu\right), \left(\partial_\mu + ig\tilde{W}_\mu\right)\right] = \partial_\mu\widetilde{W}_\nu - \partial_\nu\tilde{W}_\mu + ig\left[W_\mu, W_\nu\right], \tag{7.27}$$

$$W^i_{\mu\nu} = \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon^{ijk}W^j_\mu W^k_\nu \qquad (i,j,k = 1,2,3). \tag{7.28}$$

$B_{\mu\nu}$ remains invariant under gauge transformation, while $\widetilde{W}_{\mu\nu}$ transforms covariantly ($\widetilde{W'}_{\mu\nu} = U_L\widetilde{W}_{\mu\nu}U_L^\dagger$).

The Lagrangian density for electroweak interaction is given by

$$\mathcal{L}_0 = \mathcal{L}_0 + \mathcal{L}_{\text{KIN}} \tag{7.29}$$

where

$$\begin{aligned} \mathcal{L}_0 &= \textstyle\sum_{j=1}^3 i\bar{\psi}_j(x)\gamma^\mu D_\mu\psi_j(x), \\ \mathcal{L}_{\text{KIN}} &= -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{2}\text{Tr}\left[\widetilde{W}_{\mu\nu}\widetilde{W}^{\mu\nu}\right] = -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}W^i_{\mu\nu}W_i^{\mu\nu}. \end{aligned} \tag{7.30}$$

Since the field strengths $W^i_{\mu\nu}$ contain a quadratic piece, the Lagrangian density $\mathcal{L}_{\text{KIN}}$ gives rise to cubic and quartic self-interactions among the gauge fields. The strength of these interactions is given by the same $SU(2)_L$ coupling g which appears in the fermionic piece of the Lagrangian.

The gauge symmetry forbids the writing of a mass term for the gauge bosons. Fermionic masses are also not possible, because the would communicate the left- and right-handed fields, which are different transformation properties, and therefore would produce an explicit breaking of the gauge symmetry.

7.5 Spontaneus symmetry breaking

When a theory es symmetric with respect to a symmetry group, but requires that one element of the group is distinct, then *spontaneus symmetry braking* has occurred. In order to generate masses, we need to break the gauge symmetry in some way; however, we also need a fully symmetric Lagrangian to preserve renormalizability. This dilemma may be solved by the possibility of getting non-symmetric results from an invariant Lagrangian. Let us considerer a Lagrangian, which

1. Is invariant under a group G of transformations.

2. Has a degenerate set of states with minimal energy, which transforms under G as the members of a given multiplet.

If one of those states is arbitrarily selected as the ground state of the system, the symmetry is said to be spontaneously broken.

Let us consider a complex scalar field $\phi(x)$, with Lagrangian density (in this example μ^2 is only a parameter and can be positive or negative)

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - V(\phi), \qquad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2 . \tag{7.31}$$

$\mathcal{L}$ is invariant under the global gauge transformation

$$\phi(x) \to \phi'(x) = \exp(i\theta)\, \phi(x) \qquad (\theta \text{ const}). \tag{7.32}$$

The ground state is obtained by minimising the potential V. We have

$$\frac{\partial V}{\partial \phi} = \mu^2 \phi^\dagger + 2\lambda \phi^\dagger (\phi^\dagger \phi) \tag{7.33}$$

so that when $\mu^2 > 0$, the minimum occurs at $\phi^\dagger = \phi = 0$. If $\mu^2 < 0$, however, there is a local maximun at $\phi = 0$, and a minimun at

$$|\phi_0| = \sqrt{-\frac{\mu^2}{2\lambda}}\,, \qquad V(\phi_0) = -\frac{\mu^2}{4\lambda} \tag{7.34}$$

In the quantum field theory this condition refers to the vacuum expectation value of ϕ

$$|\langle 0|\phi|0\rangle|^2 = -\frac{\mu^2}{2\lambda} \tag{7.35}$$

Owing to the $U(1)$ phase-invariance of the Lagrangian density, there is an infinite number of degenerate states of minimun energy (vacua) related to each other by rotation, $\phi_0(x) = |\phi_0| \exp\{i\theta\}$. By choosing a particular solution as the groun state, the symmetry gets spontaneously broken. If we parametrize the excitations over the ground state as

$$\phi(x) = |\phi_0| + \frac{\phi_1(x) + i\phi_2(x)}{\sqrt{2}}, \tag{7.36}$$

so that $\langle 0|\phi_1|0\rangle = \langle 0|\phi_1|0\rangle = 0$ (ϕ_1 and ϕ_2 are real fields). The Lagrangian density takes the form

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi_1 \partial^\mu \phi_1 + \frac{1}{2}\partial_\mu \phi_2 \partial^\mu \phi_2 + \mu^2 \phi_1^2 - \sqrt{\frac{-\mu^2}{\lambda}}\,\phi_1 \left(\phi_1^2 + \phi_2^2\right) - \frac{\lambda}{4}\left(\phi_1^2 + \phi_2^2\right)^2 . \tag{7.37}$$

Thus, ϕ_1 describes a massive state of mass $m_1^2 = -2\mu^2$, while ϕ_2 is massless (known as a *Goldstone boson*). The fact that there massless excitations associated with the spontaneous symmetry breaking (SSB) mechanism is a completely general, result, known as the Goldstone theorem.

7.6 Higgs mechanism

Gauge theories do not obey Lorentz invariance and Hilbert space with positive-definite scalar products simultaneosly. In a covariant gauge, the theory contains states of negative norm. In a gauge in hich the theory has only states of positive norm it is not manifestly covariant. In consequence, Goldstone's theorem does not hold and the so called Higgs mechanism operates.

There is, at present, impressive experimental evidence for the electroweak gauge theory with the gauge symmetry spontaneously broken. In this theory, electromagnetism and the weak force are combined in a non-trivial way. The $SU(2) \times U(1)$ group is the minimal one which contains the electromagnetic and weak currents.

Let us consider un $SU(2)_L$ doublet of complex scalar fields

$$\phi(x) \equiv \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} \tag{7.38}$$

The gauged scalar Lagrangian of the Goldstone model in Eq.(7.31),

$$\mathcal{L}_S = (D_\mu \phi)^\dagger D^\mu \phi - \mu^2 \phi^\dagger \phi - \lambda \left(\phi^\dagger \phi\right)^2 \qquad (\lambda > 0, \quad \mu^2 < 0), \tag{7.39}$$

where

$$D_\mu \phi = \left[\partial_\mu - ig\widetilde{W}_\mu - ig' y_\phi B_\mu\right] \phi \qquad y_\phi = Q_\phi - T_3 = \frac{1}{2}, \tag{7.40}$$

is invariant under local $SU(2)_L \otimes U(1)_Y$ transformations. $y = Q - T_3$ is the Gell-Mann-Nishijima formula (T_3 is the eigenvalue of the isospin operator T_3, Q is the electric charge and y the hypercharge) and the valor of the scalar hypercharge is fixed by the requirement of having the correct couplings between $\phi(x)$ and $A^\mu(x)$; i.e., the photon does not couple to $\phi^{(0)}$, and $\phi^{(+)}$ has the right electric charge .

The potential is very similar to the one considered before. There is a infinite set of degenerate states with minimun energy, satisfying

$$|\langle 0|\phi_2|0\rangle| = \sqrt{-\frac{\mu^2}{2\lambda}} = a \tag{7.41}$$

Once we choose a particular ground state, the $SU(2)_L \otimes U(1)_Y$ symmetry gets spontaneously broken to the electromagnetic subgroup $U(1)_{QED}$, which by construction still remains a true symmmetry of the vacuum. Acording to the Goldstone theorem three massless states should then appear.

Now, let us parametrize the scalar doublet in the form

$$\phi(x) = \exp\left\{i\frac{\sigma_i}{2}\theta^i(x)\right\} \begin{pmatrix} 0 \\ a + \dfrac{H(x)}{\sqrt{2}} \end{pmatrix} \tag{7.42}$$

with four real fields $\theta^i(x)$ and $H(x)$. The crucial point is that the local $SU(2)_L$ invariance of the Lagrangian allows us to rotate away any dependence on $\theta^i(x)$. These three fields are precisely the would-be massless Goldstone bosons associated with the SSB mechanism.

The covariant derivative (7.40) couples the scalar multiplet to the $SU(2)_L \otimes U(1)_Y$ gauge bosons. If one takes the physical (unitary) gauge $\theta^i = 0$, the kinetic piece of the Lagrangian density (7.39)

takes the form

$$(D_\mu\phi)^\dagger D^\mu\phi = \frac{1}{2}\partial_\mu H\partial^\mu H + \frac{g^2a^2}{4}\left[(W_\mu^1)^2 + (W_\mu^2)^2\right] + \frac{a^2}{4}\left(gW_\mu^3 - g'B_\mu\right)^2 + \text{cubic and quartic terms} \tag{7.43}$$

Now we define

$$Z_\mu = \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2+g'^2}} \equiv \cos\theta_W W_\mu^3 - \sin\theta_W B_\mu \tag{7.44}$$

and the orthogonal field

$$A_\mu = \frac{g'W_\mu^3 + g'B_\mu}{\sqrt{g^2+g'^2}} \equiv \sin\theta_W W_\mu^3 + \cos\theta_W B_\mu \tag{7.45}$$

where the *Weinberg angle* θ_W is given by

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \qquad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}}. \tag{7.46}$$

The vacuum expectation values of the neutral scalar has generated a quadratic term for $W^\pm$ and the Z, i.e., those bosons have acquired masses:

$$M_Z\cos\theta_W = M_{W^1} = M_{W^2} = \frac{ag}{\sqrt{2}} \tag{7.47}$$

and A_μ is massless. A_μ is identified with the electromagnetic field and the charged doblet of massive vector particles $W_\mu^\pm$ are related to W_μ^1 and W_μ^2 by

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2). \tag{7.48}$$

7.6.1 The Higgs boson

The scalar Lagrangian in Eq.(7.43) has introduced a new scalar particle into the model: the Higgs H. In terms of the physical fields (unitary gauge), takes the form

$$\mathcal{L}_S = \frac{1}{8}M_H^2v^2 + \mathcal{L}_H + \mathcal{L}_{HG^2}, \tag{7.49}$$

where

$$\mathcal{L}_H = \frac{1}{2}\partial_\mu H\partial^\mu H - \frac{1}{2}M_H^2H^2 - \frac{M_H^2}{2v}H^3 - \frac{M_H^2}{8v^2}H^4, \tag{7.50}$$

$$\mathcal{L}_{HG^2} = \left(M_W^2W_\mu^\dagger W^\mu + M^2Z_\mu Z^\mu\right)\left(1 + \frac{2}{v}H + \frac{H^2}{v^2}\right) \tag{7.51}$$

and the Higgs mass is given by $M_H = \sqrt{-2\mu^2}$. All Higgs couplings are determined by M_H, M_W, M_Z and the vacuum expectation value v.

Chapter 8

Some topics in modern gauge theory

8.1 Lattice gauge theory

8.1.1 Introduction

The perturbative approach to field theory has led to impressive results in weakly interacting theories. For example, the anomalous magnetic moment of electron derived from QED is the quantitatively best understood quantity in physics. Still, even at weak coupling the perturbative approach to field theory is not entirely satisfactory. And in QCD the perturbative regularization is completely useless at low energies.

Confinement or the Higgs mechanism are non-perturbative phenomena. In order to study them from first principles one must first define the theory beyond perturbation theory. The lattice regularization provides a clean way of doing this by replacing the space-time continuum with a discrete mesh of lattice points. In order to recover the continuum limit, the theory must be renormalized by sending the lattice spacing to zero while adjusting the bare coupling constants appropriately. In fact, the lattice is a beautiful regularization because it is local and it respects local gauge symmetries. The fact that it violates some space-time symmetries is less important, because these symmetries are automatically recovered in the continuum limit.

8.1.2 Lattice Yang-Mills theory

Let us begin by defining the simplest lattice in for dimensions, a Euclidean hypercubical lattice with equal lattice spacing a in the x, y, z, and t direction. If we take the limit as $a \to 0$, the our action should reduce to the usual Yang-Mills action.

Wegner and Wilson, as well as Smith, independently introduced the concept of a *parallel transporter* $U_{x,\hat{\mu}} \in SU(N)$ connecting neighboring lattice points x and $x + \hat{\mu}$, where $\hat{\mu}$ defines a direction in the μth lattice direction. The parallel transporter is related to an underlying continuum gauge field $A_\mu(x) = igA^a_\mu(x)T^a$ by

$$U_{x,\mu} = \mathcal{P}\exp \int_0^a dt\, A_\mu(x + \hat{\mu}t). \tag{8.1}$$

where $\mathcal{P}$ denotes path-ordering. Under a non-abelian gauge transformation the parallel transporter transforms as

$$U'_{x,\mu} = \Omega_x U_{x,\mu} \Omega^\dagger_{x+\hat{\mu}} \tag{8.2}$$

We define a *plaquette* as a square face of the lattice with dimensions $a \times a$. Wilson has constructed the Yang-Mills action by multiplying parallel transporters around an elementary plaquette. The standard Wilson action is constructed as a sum over all plaquettes

$$S_{YM}[U] = -a^4 \sum_{x,\mu,\nu} \frac{1}{g^2 a^2} Tr[U_{x,\mu} U_{x+\hat{\mu},\nu} U^\dagger_{x+\hat{\nu},\nu} U^\dagger_{x,\nu} + U_{x,\nu} U_{x+\hat{\nu},\mu} U^\dagger_{x+\hat{\mu},\nu} U^\dagger_{x,\mu}]. \tag{8.3}$$

It reduces to the continuum Yang-Mills action in the limit $a \to 0$.

To fully define the path integral we must also consider the measure. The lattice functional integral is obtained as an integral over all configurations of parallel transporters $U_{x,\mu}$, i.e.

$$Z = \prod_{x,\mu} \int_{SU(N)} dU_{x,\mu} \exp\left(-S_{YM}[U]\right). \tag{8.4}$$

One integrates independently over all link variables using the local Haar measure $dU_{\mu,x}$ for each parallel transporter. The Haar measure is a left (and right) invariant measure, i.e.

$$\int_{SU(N)} dU\, f(\Omega U) = \int_{SU(N)} dU\, f(U\Omega) = \int_{SU(N)} dU\, f(U), \tag{8.5}$$

for any function $f(U)$ and for any $SU(N)$ matrix Ω. It is convenient to normalize the measure such that

$$\int_{SU(N)} dU = 1. \tag{8.6}$$

For compact groups like $SU(N)$ the integration is over a finite domain. This makes it unnecessary to fix the gauge in lattice Yang-Mills theory because the functional integral is finite even without gauge fixing. This is another important advantage of the formulation using parallel transporters.

8.1.3 Scalars and Fermions on the lattice

We now generalize these results to put scalars and fermions on the lattice.

To put scalars on the lattice, we must make the substitution

$$\partial_\mu \phi \to \frac{\phi_{x+\hat{\mu}} - \phi_x}{a} \tag{8.7}$$

With this simple substitution, we find that the scalar action becomes

$$S = \int d^4x \left(\frac{1}{2}\partial_\mu \phi \partial^\mu \phi + \frac{1}{2} m^2 \phi^2 + \frac{\lambda}{4!}\phi^4\right) \tag{8.8}$$

$$\to \sum_x \left[\frac{a^2}{2} \sum_{\mu=1}^{4} (\phi_{x+\hat{\mu}} - \phi_x)^2 + a^4 \left(\frac{m^2}{2}\phi_x^2 + \frac{\lambda}{4!}\phi_x^4\right)\right] \tag{8.9}$$

To calculate the propagator of the scalar particle on the lattice, we will find it convenient to go to momentum space. We wish to replace ϕ_n with its Fourier transform $\phi(k)$. We will define

$$\phi_x = \int \frac{d^4k}{(2\pi)^4} e^{ik\cdot x} \phi(k) \tag{8.10}$$

We will arbitrarily truncate the integral, since wavelengths smaller than twicw the size of the lattice can be discarded. We will take

$$-\frac{\pi}{a} \leq k_\mu \leq \frac{\pi}{a} \tag{8.11}$$

Now let us insert the Fourier expansion of ϕ_x onto the free action of the scalar field on the lattice. The free part can be calculated by taking a double integral over k and k'

$$a^4 \sum_x \int \frac{d^4k}{(2\pi)^4} \int \frac{d^4k'}{(2\pi)^4} e^{i(k+k')\cdot x}(e^{iak'_\mu} - 1)(e^{iak'_\mu} - 1) \tag{8.12}$$

$$= \int \frac{d^4k}{(2\pi)^4}(e^{iak'_\mu} - 1)(e^{-iak'_\mu} - 1) = 4\int \frac{d^4k}{(2\pi)^4} \sin^2(ak_\mu/2) \tag{8.13}$$

Inserting this back into the free action, we now have:

$$S = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4} \left(\sum_\mu \frac{4}{a^2} \sin^2(ak_\mu/2) + m^2 \right) \phi(-k)\phi(k) \tag{8.14}$$

This differs from the usual propagator defined in momentum space. Normally, the Euclidean Klein-Gordon equation has a propagator given by $1/(k^2 + m^2)$. On the lattice, the propagator is generated by taking the inverse of

$$k^2 + m^2 \to \sum_\mu \frac{4}{a^2} \sin^2(ak_\mu/2) + m^2 \tag{8.15}$$

In the limit as $a \to 0$, we find that two expressions are identical. For large k, the two expressions differ noticeably. However, large values of k are cut off.

When we try to place fermions on the lattice we find difficulties. As before, we make the substitution

$$\partial_\mu \psi \to \frac{\psi_{x+\hat{\mu}} - \psi_x}{a} \tag{8.16}$$

With this substitution, our lattice fermionic action becomes

$$S = \sum_x \left(\frac{a^3}{2} \sum_{\mu=1}^{4} \left(\bar{\psi}_x \gamma_\mu U_{x,\mu} \psi_{x+\hat{\mu}} - \bar{\psi}_{x+\hat{\mu}} \gamma_\mu U^\dagger_{x,\mu} \psi_x \right) + m a^4 \bar{\psi}_x \psi_x \right). \tag{8.17}$$

As before, we take the Fourier transform of the ψ_x field. This gives us the action

$$S = \int \frac{d^4k}{(2\pi)^4} \bar{\psi}(-k) \left(i \sum_\mu \gamma^\mu \frac{\sin(ak_\mu)}{a} + m \right) \psi(k). \tag{8.18}$$

Unfortunately, the lattice fermion theory does not give us the correct continuum limit because the lattice fermion propagator has three minima ($k = 0$, $K = \pm\pi/a$). One solution to the problem is introducing the Wilson term

$$a^4 \sum_x \frac{1}{2a} \sum_{\mu=1}^{4} \left(2\bar{\psi}_x \psi_x - \bar{\psi}_x U_{x,\mu} \psi_{x+\hat{\mu}} - \bar{\psi}_{x+\hat{\mu}} U^\dagger_{x,\mu} \psi_x \right) \tag{8.19}$$

If we now calculate the momentum-space contribution of this term and add it to the previous one, we find

$$S = \int \frac{d^4k}{(2\pi)^4} \bar{\psi}(-k) \left(i \sum_\mu \gamma^\mu \frac{\sin(ak_\mu)}{a} + m - \sum_\mu \frac{\cos(ak_\mu) - 1}{a} \right) \psi(k). \tag{8.20}$$

The second term, containing the cosine, preserves the original minimun at $k = 0$ but eliminates the unwanted one.

8.1.4 Confinement in the strong coupling limit of QCD

In lattice gauge theory it is straightforward to prove confinement for large values of the bare gauge coupling g. In the strong coupling region, however, the We define the parameter

$$W_\mathcal{C} = \mathrm{Tr} \prod_{(x,\mu} \in \mathcal{C}\, U_{x,\mu}. \tag{8.21}$$

where we take the product around a discretezed loop $\mathcal{C}$. We will be interested in the behaviour of $W_\mathcal{C}$ where $\mathcal{C}$ is a rectangular loop with width R in one spatial direction and length T in the time direction, in the limit of large T.

Assuming that there is no phase transition between the strong and weak coupling regins, the derivation of confinement in the strong coupling regime would carry over to the continuum limit. In the strong coupling expansion we expand in powers og $1/g$ around $g = \infty$. To leading order the pure gluon action is then simply zero. The Wilson loop operator takes the form

$$W_\mathcal{C} = U_{1ij} U_{2jk} U_{3kl} \cdots U_{Nmi}, \tag{8.22}$$

where $N = 2(R+T)$ is the number of links along the loop. Using the group integration formula

$$\int dU\, U_{ij} = 0, \quad \int dU\, U_{ij} U_{kl} = \frac{1}{N} \delta_{jk} \delta_{il}, \tag{8.23}$$

then immediatly implies $W_\mathcal{C} = 0$. By expanding the Boltzmann factor of the action to lowest order in $1/g$ we have

$$\langle W_\mathcal{C} \rangle = \frac{1}{(g^2)^{RT}}. \tag{8.24}$$

The Wilson loop is related to the static quark-antiquark potential $V(R)$ by

$$\lim_{T\to\infty} \langle W_\mathcal{C} \rangle \sim \exp(-V(R)T). \tag{8.25}$$

In QCD we expect quarks and anti-quarks to be confined to one another by a potential rising linearly at large separations R, i.e.

$$\lim_{R\to\infty} V(R) \sim \sigma R, \tag{8.26}$$

where σ is the string tension. In a confinement phase the Wilson loop hence shows an area law

$$\lim_{R,T\to\infty} \langle W_\mathcal{C} \rangle \sim \exp(-\sigma RT). \tag{8.27}$$

Confinement is indeed verified very accurately in numerical simulations of lattice Yang-Mills theories.

8.2 Supersymmetry

8.2.1 Introduction

Supersymmetry (often abbreviated as *susy*) is symmetry that relates elementary particles of one spin to other particles that differ by half a unit of spin and are known as *superpartners*. In a theory with unbroken supersymmetry, for every type of boson (integer spin) there exists a corresponding type of fermion (half odd integer) with the same mass and internal quantum numbers, and vice-versa.

There is no direct evidence for the existence of supersymmetry. It is motivated by possible solutions to several theorical problems. Since the superpartners of the Standard Model (SM) particles have not been observed, supersymmetry, if it exists, must be a broken symmetry, allowing the superparticles to be havier than the corresponding SM particles.

To have any chance at all of realizing supersymmetry, even spontaneously broken, the particle spectrum of the Standard Model must be extended. The minimal way of doing this, via the introduction of the smallest number of supplementary particles, is called the *Minimal Supersymmetric Standard Model*. Basically, for each currently known or *non-supersymmetric* particle, we supply a hypothetical *supersymmetric partner*. There is one exception: supersymmetry requires at least two Higgs fields.

The full significance of supersymmetry really emerges only when the principle of supersymmetry is considered in conjunction with the principle of the unification of the strong, electro-magnetic and weak forces. The smallest such group which contains the gauge group of the SM as a subgroup is $SU(5)$. To incorporate the gravity we need larger groups. And the inclusion of a massive neutrino can be realized only in the $SO(10)$ group or larger.

8.2.2 Superspace and superfields

We introduce four antisymmetric coordinates θ_α that form the superpartner of the usual space-time coordinate:

$$\{x^\mu, \theta_\alpha\}\,. \tag{8.28}$$

The coordinates x^μ and θ^α parametrize the *superspace*.

Superymmetry, acting on the superspace coordinates, makes the transformation

$$\begin{aligned} x^\mu &\to x^\mu + i\bar{\epsilon}\gamma^\mu\theta \\ \theta_\alpha &\to \theta_\alpha + \epsilon_\alpha \end{aligned} \tag{8.29}$$

In practice, the use of complex Dirac spinors leads to reducible representations of supersymmetry. In order to find irreducible representations, we will find it more convenient to use Weyl spinors. We will therefore split the four-component spinor into two smaller spinors

$$\theta_\alpha \equiv \begin{pmatrix} \theta^a \\ \bar{\theta}_{\dot{a}} \end{pmatrix} \qquad a = 1,2 \quad \dot{a} = 1,2. \tag{8.30}$$

where $(\theta^a)^* = \bar{\theta}^{\dot{a}}$. In this formalism, we will take a modified Weyl representation of the Dirac matrices

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}; \qquad \gamma_5 = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \tag{8.31}$$

where

$$\sigma^\mu = (\mathbf{1}, \sigma^i); \qquad \bar{\sigma}^\mu = (\mathbf{1}, -\sigma^i). \tag{8.32}$$

Then the typical spinor breaks up

$$\psi = \begin{pmatrix} \phi_a \\ \\ \bar{\chi}^{\dot{a}} \end{pmatrix}; \qquad \bar{\psi} = \psi^\dagger \gamma^0 = (\chi^a \quad \bar{\phi}_{\dot{a}}). \tag{8.33}$$

In two-spinor notation, the supersymmetric transformation in the superspace is written as

$$\begin{aligned} &x^\mu \to x^\mu + i\epsilon\sigma^\mu - i\theta\sigma^\mu\bar{\epsilon} \\ &\theta^a \to \theta^a + \epsilon^a \\ &\bar{\theta}_{\dot{a}} \to \bar{\theta}_{\dot{a}} + \bar{\epsilon}_{\dot{a}} \end{aligned} \tag{8.34}$$

A *superfield* $S(x, \theta, \bar{\theta})$ is defined as a general function of the supespace. By power spanding $S(x, \theta, \bar{\theta})$ in a power series in θ and $\bar{\theta}$, we find that the series terminates after reaching the fourth power of the spinor because of its Grassmann nature,

$$\begin{aligned} S(x, \theta, \bar{\theta}) =& \phi(x) + \theta\psi(x) + \bar{\theta}\bar{\chi}(x) + \bar{\theta}\bar{\sigma}^\mu\theta A_\mu(x) + \theta\theta f(x) + \bar{\theta}\bar{\theta}g^*(x) \\ &+ i\theta\theta\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}\bar{\theta}\theta\rho(x) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x). \end{aligned} \tag{8.35}$$

Naturally, in quantum field theory, one restricts to superfields that are either bosonic or fermionic, so that the has definite commutation and anticommutations relations with θ and $\bar{\theta}$,

$$\begin{aligned} \text{bosonic superfield} \qquad & [S, \theta^a] = [S, \bar{\theta}_{\dot{a}}] = 0 \\ \text{fermionic superfield} \qquad & \{S, \theta^a\} = \{S, \bar{\theta}_{\dot{a}}\} = 0. \end{aligned} \tag{8.36}$$

Thus, if S is bosonic, the component fields ϕ, A_μ, f, g and $D(x)$ are bosonic as well, while the fields ψ, χ, λ and ρ are fermionic. On the other hand, if S is fermionic, the component fields ϕ, A_μ, f, g and $D(x)$ are fermionic as well, while the fields ψ, χ, λ and ρ are bosonics.

On superfields, supersymmetry transformations are naturally realized in a linear way via super-differential operators (just as on ordinary fields, translations and Lorentz transformations are realized in a linear way via differential operators).

Superderivatives are defined by

$$D_a \equiv \frac{\partial}{\partial\theta^a} + i\sigma^\mu_{a\dot{a}}\bar{\theta}^{\dot{a}}\partial_\mu \qquad \bar{D}_{\dot{a}} \equiv -\frac{\partial}{\partial\bar{\theta}^{\dot{a}}} - i\theta^a\sigma^\mu_{a\dot{a}}\partial_\mu \tag{8.37}$$

where differentiation and integration of θ coordinates are defined by

$$\frac{\partial}{\partial\theta^a}(1, \theta^b, \bar{\theta}^{\dot{b}}) \equiv \int d\theta^a (1, \theta^b, \bar{\theta}^{\dot{b}}) \equiv (0, \delta^b_a, 0). \tag{8.38}$$

We also define Lorentz scalar differentials by

$$d^2\theta \equiv \frac{1}{4}d\theta_a d\theta^a, \qquad d^2\bar{\theta} \equiv \frac{1}{4}d\bar{\theta}^{\dot{a}}d\bar{\theta}_{\dot{a}}, \qquad d^4\theta \equiv d^2\theta d^2\bar{\theta}, \tag{8.39}$$

Under a supersymmetric transformations, superfield transforms as

$$\delta_\epsilon S = (\epsilon Q + \bar{\epsilon}\bar{Q})S \tag{8.40}$$

with the supercharges defined by

$$Q_a \equiv \frac{\partial}{\partial\theta^a} - i\sigma^\mu_{a\dot{a}}\bar{\theta}^{\dot{a}}\partial_\mu \qquad \bar{Q}_{\dot{a}} \equiv -\frac{\partial}{\partial\bar{\theta}^{\dot{a}}} + i\theta^a\sigma^\mu_{a\dot{a}}\partial_\mu. \tag{8.41}$$

The super-differential operators D_a and Q_a, differ only by a sign change, and generate left and right actions of supersymmetry respectively. Their relevant structure relations are

$$\{Q_a, \bar{Q}_{\dot{b}}\} = 2\sigma^\mu_{a\dot{b}}P_\mu, \qquad \{D_a, \bar{D}_{\dot{b}}\} = -2\sigma^\mu_{a\dot{b}}P_\mu \tag{8.42}$$

where $P_\mu = i\partial_\mu$. Since left and right actions mutually commute, all 4 components of D anti-commute with all components of Q: $\{Q_a, D_b\} = \{Q_a, \bar{D}^{\dot{b}}\} = 0$, and their complex conjugate relations.

The type of superfield introduced above is in general highly reducible, and the irreducible components may be found by imposing supersymmetric conditions on the superfield.

(a) The *Chiral Superfield* Φ is obtained by imposing the condition

$$\bar{D}_{\dot{a}}\Phi = 0. \tag{8.43}$$

The *antichiral superfield* $\Phi^\dagger$ is obtained by imposing $D_a\Phi^\dagger = 0$. These conditions are invariant under supersymmetry transformations of (8.40) since D or $\bar{D}$ and Q or $\bar{Q}$ anti-commute. Equation (8.43) may be solved in terms of the composite coordinates

$$x^\mu_\pm = x^\mu \pm \theta\sigma^\mu\bar{\theta}, \tag{8.44}$$

which satisfy

$$\bar{D}_{\dot{a}}\, x^\mu_+ = 0, \qquad D_a\, x^\mu_- = 0, \tag{8.45}$$

and we have (a factor of $\sqrt{2}$ has been inserted multiplying ψ to give this field standard normalization)

$$\begin{aligned}\Phi(x,\theta,\bar{\theta}) &= \phi(x_+) + \sqrt{2}\theta\psi(x_+) + \theta\theta F(x_+) \\ \Phi^\dagger(x,\theta,\bar{\theta}) &= \phi^*(x_-) + \sqrt{2}\bar{\theta}\bar{\psi}(x_-) + \bar{\theta}\bar{\theta}F^*(x_-)\end{aligned} \tag{8.46}$$

Te components fields ϕ and ψ are the scalar and left Weyl spinor fields of the chiral multiplet respectively. The field F has not appeared previously. The field equation for F is always algebraic, so that F is a non-dynamical or auxiliary field of the chiral multiplet.

(b) The *Vector Superfield* obtained by imposing the condition

$$V = V^\dagger \tag{8.47}$$

on a general superfield of the type (8.35). This condition sets $\chi = \psi$, $g = f$ and $\rho = \lambda$ in (8.35), and requires that the fields ϕ, A_μ and D be real. It is conventional to use a specific notation for vector superfields and it is convenient to define its expansion by

$$\begin{aligned}V(x,\theta,\bar{\theta}) =& v(x) + \theta\chi(x) + \bar{\theta}\bar{\chi}(x) + \theta\theta f^*(x) + \bar{\theta}\bar{\sigma}^\mu\theta A_\mu(x) \\ &+ i\theta\theta\bar{\theta}(\bar{\lambda}(x) + \frac{1}{2}\bar{\sigma}^\mu\partial_\mu\chi(x)) - i\bar{\theta}\bar{\theta}\theta(\lambda(x) + \frac{1}{2}\sigma^\mu\partial_\mu\bar{\chi}(x)) \\ &+ \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}(D(x) + \frac{1}{2}\partial_\mu\partial^\mu v(x))\end{aligned} \tag{8.48}$$

The *gauge superfield* is a special case of a vector superfield. On a single (abelian) vector superfield V, the reality condition $V^\dagger = V$ is preserved upon addition of a chiral superfield Λ and its complex conjugate $\Lambda^\dagger$, as follows,

$$V \to V' = V + i\Lambda - i\Lambda^\dagger. \tag{8.49}$$

Under this transformation, the component fields λ and D of V are unchanged, v, χ and f transform in a purely algebraic way,

$$\begin{aligned} v &\to v' = v + i\phi - i\phi^* \\ \chi &\to \chi' = \chi + i\sqrt{2}\psi \\ f &\to f' = f + iF \end{aligned} \tag{8.50}$$

while the field A_μ transforms as an abelian gauge field

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu(\phi + \phi^*). \tag{8.51}$$

Thus, it is natural to view (8.49) as the superfield generalization of a gauge transformation on an abelian gauge superfield V.

The non-abelian generalization of the gauge field is such that V takes in the Lie algebra $\mathfrak{g}$ of the gauge group G and the transformation (8.49) is replaced by the following non-linear gauge transformation law,

$$e^V \to e^{V'} = e^{-i\Lambda^\dagger} e^V e^{i\Lambda}. \tag{8.52}$$

which again preserves the reality condition $V^\dagger = V$, assuming that Λ is a chiral superfield transforming under the adjoint representation of the gauge algebra $\mathfrak{g}$.

As is clear from (8.50), (and an analogous result holds for the non-abelian case), the components fields v, χ, and f may be gauged away in an algebraic way, without implying any dynamical constraints. The gauge in which this is achieved is called *Wess-Zumino gauge*, and is almost always imposed when performing practical calculations in the superfield formulation. What remains is the gauge superfield in Wess-Zumino gauge, given by

$$V(x, \theta, \bar\theta) = \bar\theta\bar\sigma^\mu A_\mu(x) + i\theta\theta\bar\theta\bar\lambda(x) - i\bar\theta\bar\theta\theta\lambda(x) + \frac{1}{2}\theta\theta\bar\theta\bar\theta D(x). \tag{8.53}$$

The component fields A_μ and λ are the gauge and gaugino fields of the gauge multiplet respectively, as discussed previously. The field D has not appeared previously and is an *auxiliary field*, just as F was an auxiliary field for the chiral multiplet.

The role of the auxiliary fields F and D in the superfield formalism is to provide a linearization of the supersymmetry transformations, as well as to allow for an off-shell realization on the fields of the supersymmetry algebra, as given in (8.40).

We can generalize the supersymmetry theory including two or more supersymmetries. If $\mathcal{N}$ is the number of supersymmetries, the supercharges Q^I ($I = 1, \cdots, \mathcal{N}$) transforms as Weyl spinors of $SO(3,1)$ and are translation invariant, so that $[P_\mu, Q^I_\alpha] = 0$. The remaining super-Lie algebra structure relations are

$$\begin{aligned} \left\{Q^I_\alpha, \bar Q_{\dot\beta J}\right\} &= 2\sigma^\mu_{\alpha\dot\beta} P_\mu \delta^I_J \\ \left\{Q^I_\alpha, Q^J_\beta\right\} &= 2\epsilon_{\alpha\beta} Z^{IJ} \end{aligned} \tag{8.54}$$

Here, we have used 2-component spinor notation, which is related to 4-component Dirac spinor by

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar\sigma^\mu & 0 \end{pmatrix} \qquad Q^I = \begin{pmatrix} Q^I_\alpha \\ \bar Q_{\dot\alpha I} \end{pmatrix} \tag{8.55}$$

By construction, the generators Z^{IJ} are anti-symmetric in the indices I and J, and commute with all generators of the supersymmetry algebra. For the last reason, the Z^{IJ} are usually referred to as *central charges*, and we have

$$Z^{IJ} = -Z^{JI} \qquad [Z^{IJ}, \text{anything}] = 0. \tag{8.56}$$

Note that for $\mathcal{N}=1$, the anti-symmetry of Z implies that $Z=0$.

The supersymmetry algebra is left invariant under a global phase rotation of all supercharges Q^I_α, forming a group $U(1)_R$. In addition, when $\mathcal{N}>1$, the different supercharges may be rotated into one another under a unitary transformation, belonging to $SU(\mathcal{N})_R$. These (automorphism) symmetries of the supersymmetry algebra are called *R-symmetries*. In quantum field theories, part or all of these R-symmetries may be broken by anomaly effects.

8.2.3 $\mathcal{N}=1$ Supersymmetric Lagrangians

Lagrangians invariant under supersymmetry are just customary Lagrangians of gauge, spin 1/2 fermion and scalar fields, with certain special relations amongst the coupling constants and masses. For our purposes, the Lagrangians of interest are of two restricted kinds

1. Renormalizable $\mathcal{N}=1$ supersymmetric gauge theories;
2. More general *low energy effective* $\mathcal{N}=1$ *supersymmetric theories*, with the property that any monomial term in the Lagrangian has a total of no more than two derivatives on all boson fields and no more than one derivative on all fermion fields. Such restricted Lagrangian may be viewed as describing phenomena in the limit of low energy and momenta, and are well familiar from soft pion physics.

In 4 space-time dimensions, all Lagrangians in group (1) automatically belong in group (2). Thus, we seek to construct all Lagrangians in (2).

We consider first the case of only the $\mathcal{N}=1$ gauge multiplet (A_μ, λ_α), and proceed by writing down all possible gauge invariant polynomial terms of dimension 4 using minimal coupling. One finds

$$\mathcal{L} = -\frac{1}{2g^2}\mathrm{tr}F_{\mu\nu}F^{\mu\nu} + \frac{\theta}{8\pi^2}\mathrm{tr}F_{\mu\nu}\tilde{F}^{\mu\nu} - \frac{i}{2}\mathrm{tr}\bar{\lambda}\bar{\sigma}^\mu D_\mu\lambda, \tag{8.57}$$

where g is the gauge coupling, θ is the instanton angle, the field strength is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + i[A_\mu, A_\nu]$, $\tilde{F}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\lambda\rho}F^{\lambda\rho}$ is the Poincaré dual of F, and $D_\mu = \partial_\mu\lambda + i[A_\mu, \lambda]$. $\mathcal{L}$ is automatically invariant under the $\mathcal{N}=1$ supersymmetry transformations

$$\begin{aligned} A_\mu \to A_\mu &= i\bar{\xi}\bar{\sigma}_\mu\lambda - i\bar{\lambda}\bar{\sigma}_\mu\xi \\ \lambda \to \lambda &= \sigma^{\mu\nu}F_{\mu\nu}\xi \end{aligned} \tag{8.58}$$

where ξ is a spin 1/2 valued infinitesimal supersymmetric parameter.

Unfortunately, as soon as scalar fields are to be included, such as is the case when dealing with chiral multiplets, it is no longer so easy to guess supersymmetry invariant Lagrangians.

$$\mathcal{L} = -\partial_\mu\phi^*\partial^\mu\phi - i\bar{\psi}\bar{\sigma}^\mu\partial_\mu\psi - \left|\frac{\partial U}{\partial\phi}\right|^2 - \mathrm{Re}\left(\psi\psi\frac{\partial^2 U}{\partial\phi^2}\right) \tag{8.59}$$

where $U(\phi)$ is forced by $\mathcal{N}=1$ supersymmetry to be a complex analytic (holomorphic) scalar funtion of ϕ, called *superpotential*. Renormalizability furthermore restricts U to be a polynomial of degree no larger than 3.

8.3 Magnetic monopoles

8.3.1 The Dirac monopole

Before we discuss the properties of the gauge monopole, let us review the properties of the Dirac magnetic monopole found in ordinary electrodynamics. In order to describe a magnetic monopole within electrodynamics, Dirac (1931) used the vector potential

$$A_x = -g\frac{y}{r(z-r)}, \quad A_y = g\frac{x}{r(z+r)}, \quad A_z = 0 \tag{8.60}$$

where $r = \sqrt{x^2+y^2+z^2}$ and g is the "magnetic charge". A strange feature of the potential (8.60) is that it is singular on the positive z axis. Nevertheless, the corresponding field strength

$$F_{\mu\nu}dx^\mu \wedge dx^\nu = \frac{g}{r^3}(x\,dy \wedge dz + y\,dz \wedge dx + z\,dx \wedge dy) \tag{8.61}$$

has a singularity only at $r = 0$, which is unavoidable for a point source. Comparing with (3.6) we see that the electric field vanishes and that the magnetic induction is given by

$$\mathbf{B} = \frac{g}{r^3}\mathbf{r} = -g\nabla\left(\frac{1}{r}\right) \tag{8.62}$$

where $\mathbf{r} = (x, y, z)$. Since $\nabla^2 1/r = -4\pi\delta^3 r$, we have

$$\nabla\cdot\mathbf{B} = 4\pi g\delta^3 r \tag{8.63}$$

corresponding to a point magnetic charge.

Next we consider a gauge transformation

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu f, \tag{8.64}$$

and for convenience, we introduce polar coordinates (r, θ, φ) and get

$$A_r = A_\theta = 0, \quad A_\varphi = -g(1+\cos\theta). \tag{8.65}$$

Choosing

$$f(x) = e^{2gi\varphi} \tag{8.66}$$

we obtain

$$A'_\varphi = g(1-\cos\theta) \tag{8.67}$$

or

$$A'_x = g\frac{y}{r(z+r)}, \quad A'_y = -g\frac{x}{r(z+r)}, \quad A'_z = 0 \tag{8.68}$$

So our gauge transformation has shifted the singularity from the positive to negative z axis. Note, however, that f is singular along the whole z-axis, because there the azimutal angle φ is not defined. Furthermore, f is single-valued only if $2g$ is an integer. This singularity (called "string") can be moved around by gauge transformations and is therefore unobservable if and only if $g \in \mathbb{Z}$. This is *Dirac's quatization condition*. Furthermore, it is easy to show that Dirac quantization condition implies electric charge quantization and the electric charge (in units $\hbar = c = 1$) is given by

$$e = \frac{n}{2g}, \tag{8.69}$$

where n is an integer number.

8.3.2 The 't Hooft-Polyakov monopole

In the context of Maxwell's electrodynamics, with abelian gauge group $U(1)$, it is clear that althoug magnetic charges may be "added "to the theory, there is no necessity for doing this. A theory with monopoles is more symmetric between electricity and magnetism that one without, but this does not amount to a requirement that monopoles exist. However, when the gauge symmetry is enlarged to a non-abelian group and spontaneous symmetry breaking is introduced, the field equations yield a solution which corresponds to a monopole. The theoretical posibility of monopoles of this type was discovered in 1974 by 't Hooft and Polyakov.

We considerer a theory with an $SO(3)$ symmetry group, containing the gauge field A^a_μ and an isotopic vector Higgs field ϕ^a. The Lagrangian density is given by

$$\mathcal{L} = -\frac{1}{4}F^a_{\mu\nu}F^{a,\mu\nu} + \frac{1}{2}\left(D_\mu\phi^a\right)\left(D^\mu\phi^a\right) - \frac{m^2}{2}\phi^a\phi^a - \lambda\left(\phi^a\phi^a\right)^2 \tag{8.70}$$

where

$$\begin{aligned} F^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + e\epsilon^{abc}A^b_\mu A^c_\nu, \\ D_\mu\phi^a &= \partial_\mu\phi^a + e\epsilon^{abc}A^b_\mu\phi^c. \end{aligned} \tag{8.71}$$

We are interested in static solutions in which the gauge potentials have the non-trivial form

$$A^a_i = \frac{1-h(r)}{er^2}\sum_{a,i,j=1}^{3} T_a\varepsilon_{aij}x^j A^a_0 = 0 \tag{8.72}$$

and the scalar field has the form

$$\phi^a = \frac{x^a}{r^2}f(r), \tag{8.73}$$

where $a, i, j = 1, 2, 3$. These expressions have a remarkable form because of the mixing they employ between space and isospin indices. The field equations imply the following system of ordinary non-linear equations for f and h:

$$\begin{aligned} r^2\frac{d^2f}{dr^2} &= f\left(2h^2 - m^2r^2 + 4\lambda f^2\right), \\ r^2\frac{d^2h}{dr^2} &= h\left(h^2 - 1 + f^2\right). \end{aligned} \tag{8.74}$$

The boundary conditions are

$$\begin{aligned} &h(r) \to 0, \quad f(r)/r \to \beta = \frac{m}{2\sqrt{\lambda}} \quad \text{as} \quad r \to \infty, \\ &h(0) = 1, \quad f(0) = 0, \end{aligned} \tag{8.75}$$

and one finds

$$h(r) \to c_1 re^{-\beta r}, f(r) \to \beta r + c_2 e^{-m\sqrt{2}r} \tag{8.76}$$

as $r \to \infty$, whereas for $r \to 0$

$$\begin{aligned} h(r) &\to 1 + c_3 r^2, \\ f(r) &\to c_4 r^2. \end{aligned} \tag{8.77}$$

Here $c_1, \ldots, c_4$ are constants, which may depend on m and λ. In particular, (8.77) shows that the fields A and ϕ are not singular at the origin. The asymptotic forms of the gauge and scalar fields constituting a 't Hooft-Polyakov monopole.

How is the 't Hooft-Polyakov monopole related to the Dirac monopole? To answer this question we perform a gauge transformation

$$\begin{aligned} A_\mu &\to A' = SA_\mu S^{-1} + S\partial_\mu S^{-1} \\ \phi &\to S\phi S^{-1} \end{aligned} \tag{8.78}$$

with an $SO(3)$-valued function S. We choose

$$S = R \begin{pmatrix} \cos\frac{\theta}{2} & \mathrm{e}^{-i\varphi}\sin\frac{\theta}{2} \\ -\mathrm{e}^{i\varphi}\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}. \tag{8.79}$$

Here θ, φ are polar coordinates and R is the homomorphism $SU(2) \to SO(3)$. So we rotate the scalar field, wich initially points the radial direction, such that it becomes parallel to the x^3-axis:

$$\begin{aligned} \phi'^a &= \delta_{a3} f(r)/r \\ A'^1 &= -h(r)(\sin\varphi d\theta + \cos\varphi \sin\theta d\varphi), \\ A'^2 &= h(r)(\cos\varphi d\theta - \sin\varphi \sin\theta d\varphi), \\ A'^3 &= -2(1-\cos\theta)d\varphi. \end{aligned} \tag{8.80}$$

Note that S and A'^3 are singular along the negative z-axis. As $r \to \infty$, h vanishes exponentially, and only the third component of the gauge field survives. Comparing this potential with potential of the Dirac monopole we conclude that 't Hooft-Polyakov monopole seen from afar looks like a Dirac with twice magnetic charge unit.

Chapter 9

Mathematical foundations of the Quantum Field Theory

9.1 Introduction

The success of relativistic quantum field theory calculations and of perturbative renormalization also led a logical puzzle: is there any phisically-relevant, relativistic quantum field theory that is also mathematically consistent? Put differently, can one give a mathematically complete example of any non-linear theory, relevant for the description of interacting particles, whose solutions incorporate relativistic covariance, positive energy and causality? In four dimensions, this has focused attention on finding a solution to a non-abelian or Yang-Mills theory. Assuming a positive answer to this existence question, can one develop a calculational scheme to determine properties of such an example, both perturbatively and non-perturbatively? Strong-coupling calculations, as wells as calculations near critical values of the coupling constants, have been the most elusive to understand. Thus one wants to understand both the quantive structure of field theories, as well as qualitative features such as the dependence of the theories as functions on the space of coupling constant parameters.

Since the inception of quantum field theory, two central methods have emerged to show the existence of quantum fields on non-compact configuration space (such as Minkowski space). These methods are

(a) Find an exact solution in closed form.

(b) Solve a sequence of approximate problems, and stablish convergence of these solutions to desire limit.

Exact solutions may be available for non-linear fields for special values of the coupling. They might be achieved after clever changes of variables. In the case of four-dimensional Yang-Mills theory, an exact solution appears unlikely, though there might some day be asymptotic solution.

The second method is to use mathematical approximations to show the convergence of approximate solutions to exact solutions of the non-linear problems, known as *constructive quantum field theory*, or CQFT. Two principal approaches —studying quantum theory on Hilbert space, and studying classical functional integrals—are related by Osterwalder-Schrader construction.

9.2 The Wightman axioms

The *Wightman axioms* are an attempt at mathematically rigorous formulation of quantum field theory. Arthur Wightman formulated the axioms in the early 1950s but they were first published only in 1964, after Haag-Ruelle scattering theory affirmed their significance.

The axioms exist in the context of constructive quantum field theory, and they are meant to provide a basis for rigorous treatment of quantum fields, and strict foundation for the perturbative methods use. One of the *Millenium Problems* is to realize the Wightman axioms in the case of Yang-Mills fields.

One basic idea of the Wightman axioms is that there is a Hilbert space upon wich the Poincaré group acts unitarily. In this way, the concepts of energy, momentum, angular momentum and center of mass (corresponding to boost) are implemented.

There is also a stability assumption which restricts the spectrum of the four-momentum to the positive light cone (and its boundary). However, this isn't enough to implement locality. For that, the Wightman axioms have position dependent operators called *quantum fields* which form covariant representations of the Poincaré group.

Since quantum fields theory suffers from ultraviolet problems, the value of a field at a point is not well-defined. To get around this, the Wightman axioms introduce the idea of smearing over a test-function to tame the UV divergences which arise even in a free field theory. Because the axioms are dealing with unbounded operators, the domains of the operators have to be specified.

The Wightman axioms restrict the casual structure of the theory by imposing either commutativity or anticommutativity between spacelike separated fields.

They also postulate the existence of a Poincaré-invariant state called the vacuum and demand it is unique. Moreover, the axioms assume that the vacuum is *cyclic*, i.e., that the set of all vectors which can be obtained by evaluating at the vacuum state elements of the polynomial algebra generated by the smeared field operators is a dense subset of the whole Hilbert space.

Lastly, there is a primitive causality restriction which states that any polynomial in the smeared fields can be arbitrarily accurately approximated (i.e., is the limit of operators in the weak toology) by polynomials over fields smeared over test functions with support in any open subspace of Minkowski space whose casual closure is the whole Minkowski space itself.

Axioms

W0 (assumptions of relativistic quantum mechanics). An ensemble corresponding to $U(a, L)|v\rangle$ is to be interpreted with respect to the coordinates $x' = L^{-1}(x - a)$ in exactly the same way as an ensemble corresponding to $|v\rangle$ is interpreted with respect to the coordinates x; and similarly for the odd subspaces.

The group of space-time translations is commutative, and so the operators can be simultaneously diagonalised. The generators of these groups give us four self-adjoint operators, P_0, P_j, $j = 1, 2, 3$, which transform under the homogeneous group as four-vector, called the energy-momentum vector.

The second part of the zeroth axiom of Wightman is that the representation $U(a, A)$ fulfills the spectral condition that the simoultaneous spectrum of energy-momentum is contained in the forward cone:

$$P_0^2 - P_j P_j \geq 0.$$

The third part of the axiom is that there is a unique state, represented by a ray in the Hilbert

space, which is invariant under the action of the Poincaré group. It is called vacuum.

W1 (assumptions on the domain and continuity of the field). For each test function f, there exist a set of operators $A_1(f), \ldots, A_n(f)$ which, together with their adjoints, are defined on a dense subset of the Hilbert state space, containing the vacuum. The field A are operator-valued tempered distributions. The Hilbert state space is spanned by the field polynomials acting on the vacuum (cyclicity condition).

W2 (transformation law of field). The fields are covariant under the action of Poincaré group, and they transform according to some representation S of the Lorentz group, or $SL(2, C)$ if the spin is not integer:

$$U(a, L)^{\dagger} A(x) U(a, L) = S(L) A\left(L^{-1}(x - a)\right).$$

W3 (local commutativity or microscopic causality). If the supports of two fields are space-like separated, then the fields aeither commute or anticommute.

Cyclicity of a vacuum, and uniqueness of a vacuum are sometimes considered separately. Also, there is property of asymptotic completeness- that Hilbert state space is spanned by the asymptotic spaces H^{in} and H^{out}, appearing in the collision S-matrix. The other important property of field theory is *mass gap* which is not required by the axioms - that energy-momentum spectrum has a gap between zero and some positive number.

Consequences of the axioms

From these axioms, certain general theorems follow:

1. PCT theorem: there is general symmetry under change of parity, particle-antiparticle reversal and time inversion.
2. Connections between spin and statistic. Field which transform according to half integer spin anticommute, while those with integer spin commute.

The Wightman framework does not cover gauge theories. Even in abelian gauge theories conventional approaches start off with a Hilbert space (it's not a Hilbert space, but physicist call it a Hilbert space) with an indefinite norm and the physical states and physical operators belong to a cohomology. This obviously is not covered anywhere in the Wightman framework. However as shown by Schwinger, Christ and Lee, Groboc, etc., canonical quantization of gauge theories in Coulomb gauge is possible with an ordinary Hilbert space, and this might be the way to make them fall under the applicability of the axiom systematics.

Currently, there is no proof that the Wightman axioms can be satisfied for interacting theories in dimension 4. In particular, the Standard Model of particle physics has no mathematically rigorous foundations.

9.3 The Euclidean axioms

The Euclidean axioms appear simpler than the Wightman axioms. They only entail

- A regularity assumption.
- Euclidean covariance.

- Reflection positivity.
- Clustering

Osterwalder and Schrader formulated these axioms in terms of the Euclidean Schwinger functions, or Green's functions for a Euclidean field. They proved a remarkable equivalence theorem relating their axioms to the Wightman axioms.

Under certain technical assumptions, it has been shown (Osterwalder-Schrader recontruction theorem) that Euclidean QFT can be Wick-rotated into a Wightman QFT. This theorem is the key tool for the constructions of interacting theories in dimension 2 and 3 which satisfay the Wightman axioms.

The Osterwalder-Schrader axioms on Euclidean Green's functions (with a natural bound on the growth of the n^{th} Green's funtions) are equivalent to the Wightman axioms on the vacuum expectation values arising from the OS-quantization of these Green's functions, along with a growth condition on the n^{th}- vacuum expectation value. Omiting clustering axiom of OS is equivalent to omitting the unique-vacuum axiom of Wightman.

When the Euclidean expectation can be given by given by a functional integral $d\mu$, these axioms can be stated in terms of the Fourier transformation $S(f) = \int e^{i\phi} d\mu(\phi)$ of the mesure $d\mu$. One says that the mesure $d\mu(\phi)$ is Euclidean-invariant if $S\left(T(O,a)f\right) = S(f)$ for all Euclidean transformations T and all real $f \in S(R^d)$. Furthermore reflection positivity translates to the property that $S(f)$ is a function of reflection-positive type: for every choice of n real functions $f_j \in S(R^d_+)$ and complex constants c_j,

$$0 \leq \sum_{i,j=1}^{n} \bar{c}_i c_j S(f_j - \Theta f_i). \tag{9.1}$$

One also says that $S(f)$ is *regular*, if the moments $\int \phi(f)^n d\mu(\phi)$ satisfy the growth of the Green's functions required by OS Theorem . One says that $S(f)$ clusters with an exponential rate m, if for all real $f, g \in C_0^\infty(R^d)$,

$$\left|S\left(f + T(t)g\right) - S(f)S(g)\right| \leq O(1)e^{-mt}. \tag{9.2}$$

Here the constant $O(1)$ may depend on f, g, and $T(t)$ denotes the time-translation subgroup $T(I, (\vec{O}, t))$. One arrives at a very interesting probabilistic consequence of this formulation.

A Euclidean invariant, reflection-positive, regular probability measure $d\mu(\phi)$ on $S'(R^d)$ yields a scalar quantum field theory satisfaying the Wightman axioms on Minkowski space-time M^d. Clustering with exponential rate m yields a mass gap $(0, m)$ in the spectrum of H.

9.4 The Yang-Mills existence and mass gap

At the present there is no mathematical definition of quantum Yang-Mills theory in four dimensions, because of the famous problem of renormalization.

To stablish existence of four-dimensional quantum field theory with gauge group G, one should define a quantum field theory with local quantum field operators in correspondence with the gauge-invariant local polynomials in the curvature F and its covariant derivatives. Correlation functions of the quantum fiel operators should be agree at short distances with the predictions of asymptotic freedom and perturbative renormalization theory, as described in textbooks.

Since the vacuum vector Ω is Poincaré invariant, it is an eigenstate with zero energy, namely $H\Omega = 0$. The positive energy axiom asserts that in any quantum fiel theory, the spectrum of H is supported in the region $[0, \infty)$. A quantum field theory a *mass gap* if H has no spectrum in the interval $(0, \Delta)$ for some $\Delta > 0$.

The Yang-Mills Millenium Prize Problem is stated as

Prove that for any compact simple gauge group G, a non-trivial quantum Yang-Mills theory exists on $\mathbb{R}^4$ and has a mass gap $\Delta > 0$.

The most promising candidate for a non-trivial and physically-interesting field theory on Minkowski 4-space is the Yang-Mills theory with an $SU(2)$ gauge group. The Yang-Mills field F is defined in terms of a Lie-algebra valued connection A,

$$F = dA + A \wedge A. \tag{9.3}$$

The Euclidean Yang-Mills Lagrangian is $\|F\|^2$, where the squared norm includes a trace over $SU(2)$ and a integral over space-time. Perturbation theory involves the study of the interaction in powers of the non-linearity arising from $A \wedge A$, and it indicates that this Yang-Mills example is asymptotically free. The physical interaction becomes weaker at high energy, and for this reason, the objections from perturbation theory suggesting the triviality of ϕ_4^4 do no carry over to Yang-Mills interactions in four-space. Furthermore, physicist expect that this example will have a mass gap.

The approach to the problem are based on approximating continuum space-time by a lattice (section , on which we define a gauge-invariant action. Understanding the functional integration and construct Euclidean field theories

Conclusions

Historically, gauge theory had a long but confused past. Weyl's first attempt at combining electromagnetism and gravitation actually ended in disaster and was rescued only by the advent of quantum mechanics, which permitted a reinterpretation if his theory that was in accord with the experimental facts. Although the Yang-Mills equation had been discovered in 1938 by O. Klein who was studying Kaluza-Klein theories, it was promptly forgotten during World War II. It was resurrected by Yang and Mills in 1954, and independently by Shaw and Utiyama, but it was unsuitable for particle interactions because it only described massless vector particles. In that time, the strong interactions appeared to be mediated by the pseudo-scalar π-mesons rather than vector-particles. Even for the weak interactions, which had a vector-like structure, the theory required that the charged currents be accompanied by a neutral one, for which there was absolutely no experimental evidence. Thus in 1954 the Yang-Mills proposal seemed destined to become no more than a theoretical toy. Only when the vector character of the weak currents emerged in 1958 was it possible to think seriously of applying the non-abelian theory to the real world.

It is not surprising that the resolution of these problems took some time. What is atonishing is that they could be resolved at all. The first major advance was in connection with the mass problem, which was solved a decade later by the invention (or, more precisely, the importation from solid-state physics) of the spontaneus symmetry-breaking mechanism. It was followed by the resolution by Gerard 't Hooft in 1971 of the renormalization (even after symmetry breaking) problem through the use of gauge-invariant renormalization techniques, notably, dimensional regularization.

These two theoretical successes spurred a serious experimental search for the neutral weak current, and its discovery led to an extensive investigation into low-energy structure of the weak interactions. The results turned out to be in complete agreement with gauge theory, in the form of standard gauge-theory model, which was based on the Lie algebra $SU(2) \times U(1)$.

Meanwhile, the situation with regard to the strong interactions was changing with the gradual realization that the nucleons and mesons were not elementary but composite particles, built of quarks. This implied that the interactions between them were non-local and therefore unlike to be fundamental. Then, in 1974, came the dramatic discovery that non-abelian gauge theories are asymptotically free, i.e., the strength of their coupling decreases with energy. The fact that asymptotic freedom agreed with high-energy observations and that its complement, infra-red slavery, would account for quark-confinement, strongly suggested that the strong interactions were described by a non-abelian gauge theory at a more fundamental level. The problem of finding the corresponding gauge group was conveniently solved by the existence of an internal unbroken continuous symmetry group, namely, the $SU(3)$ color group of the quarks. Thus was born the present picture of the strong interactions in which the quarks interact by coupling locally to $SU(3)$ gauge fields (coloured gluons) and thereby induce the observed non-local interactions of their baryon and meson composites. The confinement of the quarks and gluons means, of

course, that it is difficult to find direct evidence for this theory, just as it would be difficult to find direct evidence for atomic theory using only molecular interactions. But the indirect evidence for $SU(3)$ gauge theory is now so strong that it is almost universally accepted.

In principle, the only difference between the gluonic interactions of the quarks and the electromagnetic interactions of the electrons and protons in atomic physics is that the gluons interact with each other because of the non-abelian nature of the group. But in practice this difference is of the utmost importance. It is the self-interaction of the gluons that is responsible for asymptotic freedom (small coupling at short distances) and, together with the long-range of the gluonic interactions, it supposed to be responsible for quark confinement (strong coupling at large distances). If the supposition is correct, it means that confinenement is due partly to the fact that gluons are massless.

The elevation of the gauge fields to the level of the gravitational field is a substancial achievement, but is by no means the end of the story. Indeed, there are two major limitations on the power of gauge theory. First, in its present form at least, it does not unify the fundamental interactions in an intrinsic way, in the sense that the coupling constants for the different interactions remais theoretically undetermined. Even in the standard electroweak model, the electromagnetic and weak coupling constants remain undetermined. Furthermore, gravitation is not unified with the other interactions. Second, gauge theories provides no answers for the questions that arise concerning the matter fields, such as the distinction between baryons and leptons, the existence of three generations of quark-lepton pairs, the origin and structure of the symmetry-breaking sector, and of the quark mixing matrix.

One way to the unification of the fundamental forces is considerate more dimensions that four. The extra dimensions can be "rolled up". For us, these rolled-up dimensions have become invisible. This idea had already been suggested by Theodor Kaluza in 1919, and was further elaborated upon by Oskar Klein. And they discovered something else. The component of the gravitational field in the direction in which space is curled up obeys exactly the same laws as Maxwell's laws of elctromagnetism. Could it be that electromagnetism is nothing but gravity in a rolled-up dimension? Einstein was enthusiastic when he heard about this idea, but it was soon realized that there is nothing one can predict with such theory, and it was abandoned. But the idea of Kaluza and Klein was rediscovered and enlarged by the experts of supergravity and superstrings. In the supersymmetric theories we enlarge the space by introducing antisymmetric coordinates.

On the other hand, we don't have a mathematically complete example of a quantum gauge theory in four-dimensional space-time, nor even a precise definition of quantum gauge theory in four dimensions. We believe that quantum field theory will have an important role in the effort to unify gravity and quantum mechanics. For mathematicians to participe in this achievement, or even to undersatand the possible results, quantum field theory must be developed as a branch of mathematics. It is important not only to understand the solution of specific problems arising from physics, but also to set such results within a new mathematical framework. One hopes that this framework will provide a unified development of several fields of mathematics and physics.

Appendix A

The Grassmann algebra

The generators η_i of an n-dimensional *Grassmann algebra* obey

$$\{\eta_i, \eta_j\} = \eta_i\eta_j + \eta_j\eta_i = 0 \tag{A.1}$$

where $i = 1, 2 \cdots, n$. In particular,

$$\eta_i^2 = 0 \tag{A.2}$$

The expansion of a function $f(\eta_i)$ only contains a finite number of terms. For example, in the 1-dimensional algebra

$$f(\eta) = a + b\eta,$$

since the quadratic term vanishes, by (A.1).

Derivatives

Because of the anticommutation relations there are two type of differentiation, left (L) and right (R)

$$\frac{\partial^L}{\partial\eta_i}(\eta_1\eta_2\cdots\eta_p) = \delta_{1i}\eta_2\cdots\eta_p - \delta_{2i}\eta_1\eta_3\cdots\eta_p + \cdots + (-1)^{p-1}\delta_{pi}\eta_1\cdots\eta_{p-1} \tag{A.4}$$

$$\frac{\partial^R}{\partial\eta_i}(\eta_1\eta_2\cdots\eta_p) = \delta_{ps}\eta_1\cdots\eta_{p-1} + \cdots + (-1)^{p-1}\delta_{1s}\eta_2\cdots\eta_p \tag{A.5}$$

It is clear that the derivative operators must obey

$$\left\{\frac{\partial}{\partial\eta_i}, \eta_j\right\} = \delta_{ij} \tag{A.6}$$

We can verefy

$$\left\{\frac{\partial}{\partial\eta_i}, \frac{\partial}{\partial\eta_j}\right\} = 0 \tag{A.7}$$

Then we have $(\partial/\partial\eta_i)^2 = 0$, which implies that there is no inverse to derivation.

Integration

We define integration by

$$\int d\eta_i = 0, \quad \int d\eta_i\, \eta_j = \delta_{ij}. \tag{A.8}$$

For a complex Grassmann variable the real and imaginary parts can be replaced by η and $\bar{\eta}$ as independent generators of Grassmann algebra.

Let η and $\bar{\eta}$ be independent Grassmann quantities. Because $\eta^2 = \bar{\eta}^2 = 0$, we have

$$e^{\bar{\eta}\eta} = 1 - \bar{\eta}\eta \tag{A.9}$$

and hence

$$\int d\bar{\eta}\, d\eta e^{\bar{\eta}\eta} = 1 \tag{A.10}$$

If η and $\bar{\eta}$ are n-dimensional vectors the result is the same.

One important result is the next formula

$$\int d\bar{\eta}\, d\eta e^{\bar{\eta}^{\mathrm{T}} A\eta} = \det A \tag{A.11}$$

To describe Fermi fields, we now make the transition to an infinite-dimensional Grassman algebra, whose generators may be denoted $C(x)$. They obey the relations

$$\begin{aligned} &\{\eta(x), \eta(y)\} = 0, \\ &\frac{\partial^{L,R}\eta(x)}{\partial\eta(y)} = \delta(x-y), \\ &\int d\eta(x) = 0; \quad \int \eta(x) d\eta(x) = 1. \end{aligned} \tag{A.12}$$

The integral (A.11) becomes functional integral over complex Grassmann variables.

Appendix B

Normal ordering and time-ordered product of operators

Normal ordering

In quantum field theory a product of quantum fields, or equivalently their creation and annihilation operators is usually said to be *normal ordered* (also called *Wick order*) when all creation operators are to the left of all annihilation operators in the product. The process of putting a product into normal order is called *normal ordering* (also called *Wick ordering*). If $\hat{O}$ denotess an arbitrary operator, then normal ordered form of $\hat{O}$ is denoted by $:\hat{O}:$ (an alternative notation is $\mathcal{N}(\hat{O})$).

Bosons

Bosons are particles which satify Bose-Einstein statistics. We will now examine the normal ordering of bosonic creation and annihilation operator products.

Single bosons

We consider the annihilation operator $\hat{a}$ and the creation operator $\hat{a}^\dagger$. These satisfy the commutator relationships

$$[\hat{a}, \hat{a}] = 0, \quad [\hat{a}^\dagger, \hat{a}^\dagger] = 0, \quad [\hat{a}, \hat{a}^\dagger] = 1, \tag{B.1}$$

$$:\hat{a}^\dagger\hat{a}: = \hat{a}^\dagger\hat{a}, \qquad :\hat{a}\hat{a}^\dagger: = \hat{a}^\dagger\hat{a}. \tag{B.2}$$

These two results can be combined with the commutation relation obeyed by $\hat{a}$ and $\hat{a}^\dagger$ to get

$$\hat{a}\hat{a}^\dagger = \hat{a}^\dagger\hat{a} + 1 =: \hat{a}\hat{a}^\dagger : +1. \tag{B.3}$$

For the exponential function

$$:\exp(\lambda\hat{a}^\dagger\hat{a}): = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\hat{a}^\dagger n\hat{a}^n. \tag{B.4}$$

Multiple bosons

For multiple bosons, $\hat{a}_i$ is the i-th boson's annihilation operator and $\hat{a}_i^\dagger$ is the i-th boson's creation operator. Here $i = 1, \ldots, n$. These satisfy the commutation relations:

$$[\hat{a}_i, \hat{a}_j] = 0, \quad [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0, \quad [\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}, \tag{B.5}$$

where $i, j = 1, \ldots, N$ and δ_{ij} denotes the Kronecker delta. For example, for three different bosons we have

$$\begin{aligned} :\hat{a}_1^\dagger\hat{a}_2\hat{a}_3: &= \hat{a}_1^\dagger\hat{a}_2\hat{a}_3 \\ :\hat{a}_2\hat{a}_1^\dagger\hat{a}_3: &= \hat{a}_1^\dagger\hat{a}_2\hat{a}_3 \\ :\hat{a}_3\hat{a}_2\hat{a}_1^\dagger: &= \hat{a}_1^\dagger\hat{a}_2\hat{a}_3 \end{aligned} . \tag{B.6}$$

Fermions

Fermions are particles which satisfy Fermi-Dirac statistics. We will now examine the normal ordering of fermionic creation and annihilation operator products.

Single fermions. We consider the annihilation operator $\hat{b}$ and the creation operator $\hat{b}^\dagger$. These satisfy the anti-commutator relationships

$$\{\hat{b}, \hat{b}\} = 0, \quad \{\hat{b}^\dagger, \hat{b}^\dagger\} = 0, \quad \{\hat{b}, \hat{b}^\dagger\} = 1. \tag{B.7}$$

To define the normal ordering of a product of fermionic creation and annihilation operators we must take into account the number of interchanges between neighbouring operators. We get a minus sign for each such interchange,

$$: \hat{b}^\dagger \hat{b} := \hat{b}^\dagger \hat{b} \quad : \hat{b}\hat{b}^\dagger := -\hat{b}^\dagger \hat{b}. \tag{B.8}$$

These can be combined, along with the anti-commutation relation, to show

$$\hat{b}\hat{b}^\dagger = 1 - \hat{b}^\dagger \hat{b} =: \hat{b}\hat{b}^\dagger : +1. \tag{B.9}$$

The normal order of any more complicated cases gives zero because there will be at least one creation or annihilation operator appearing twice. For example:

$$: \hat{b}\hat{b}^\dagger \hat{b} := -\hat{b}^\dagger \hat{b}\hat{b} = 0. \tag{B.10}$$

Multipe fermions

For multiple fermions, $\hat{b}_i$ is the i-th fermion's annihilation operator and $\hat{b}_i^\dagger$ is the i-th fermion's creation operator. Here $i = 1, \ldots, n$. These satisfy the anti-commutation relations:

$$\{\hat{b}_i, \hat{b}_j\} = 0, \quad \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0, \quad \{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij} \tag{B.11}$$

where $i, j = 1, \ldots, N$. For example, for three different fermions we have

$$\begin{array}{l} : \hat{b}_1^\dagger \hat{b}_2 \hat{b}_3 := \hat{b}_1^\dagger \hat{b}_2 \hat{b}_3 = -\hat{b}_1^\dagger \hat{b}_3 \hat{b}_2 \\ : \hat{b}_2 \hat{b}_1^\dagger \hat{b}_3 := -\hat{b}_1^\dagger \hat{b}_2 \hat{b}_3 = \hat{b}_1^\dagger \hat{b}_3 \hat{b}_2 \\ : \hat{b}_3 \hat{b}_2 \hat{b}_1^\dagger := \hat{b}_1^\dagger \hat{b}_3 \hat{b}_2 = -\hat{b}_1^\dagger \hat{b}_2 \hat{b}_3 \end{array} \tag{B.12}$$

Uses in quantum field theory

The vacuum expectation value of a normal ordered product of creation and annhilation operators is zero. This is because, denoting the vacuum state by $|0\rangle$, the creation and annihilation operators (bosonic or fermionic) satisfy

$$\langle 0|\hat{a}^\dagger = 0 \quad \text{and} \quad \hat{a}|0\rangle. \tag{B.13}$$

Any normal ordered operator therefore has a vacuum expectation value of zero. Although an operator $\hat{O}$ may satisfy $\langle 0|\hat{O}|0\rangle \neq 0$ we always have $\langle 0| : \hat{O} : |0\rangle = 0$. This is particularly useful when defining a quantum mechanical Hamiltonian. If the Hamiltonian of a theory is in normal order then the ground state energy will be zero.

Time-ordered product

In quantum field theory it is useful to take the *time-ordered* product of operators. This operation is denoted by $\mathcal{T}$. For two operators $A(x)$ and $B(y)$ that depend on spacetime locations x and y we define:

$$\mathcal{T}[A(x)B(y)] = \begin{cases} A(x)B(y) & \text{if} \quad x_0 > y_0 \\ \pm B(y)A(x) & \text{if} \quad x_0 < y_0. \end{cases} \tag{B.14}$$

Here x_0 and y_0 denote the time-coordinates of the points x and y.

Explicity we have

$$\mathcal{T}[A(x)B(y)] = \theta(x_0 - y_0)A(x)B(y) \pm \theta(y_0 - x_0)B(y)A(x), \tag{B.15}$$

where θ denotes the Heaviside step function and the $\pm$ depends if the operators are bosonic or fermionic in nature. If bosonic, then the $+$ sign is always chosen, if fermionic then the sign will depend on the number of operator interchange necessary to achieve the proper time ordering. Note that the statistical factors do not enter here.

Since the operators depend on their localization is spacetime (i.e. not just time) this time-ordering operation is only coordinate independent if operators at spacelike separated points commute. Note that the time-ordering is usually written with the argument increasing from right to left.

The S-matrix (scattering matrix) in quantum field theory is an example of a time-ordered product. The S-matrix, transforming the state $t = -\infty$ to a state $t = +\infty$, can also be thought of as a kind of "holonomy". We obtain a time-ordered expression because of the following reason:

We start with this simple formula for the exponential

$$\exp(h) = \lim_{N\to\infty} \left(1 + \frac{h}{N}\right)^N. \tag{B.16}$$

Now we consider the discretized evolution operator

$$S = \ldots (1 + h_{+3})(1 + h_{+3})(1 + h_{+2})(1 + h_{+1})(1 + h_0)(1 + h_{-1})(1 + h_{-2}) \ldots \tag{B.17}$$

where $1+h_j$ is the evolution operator over an infinitesimal time interval $[j\epsilon, (j+1)\epsilon]$. The higher order terms can be neglected in the limit $\epsilon \to 0$. The operator h_j is defined by

$$h_j = -\frac{i}{\hbar} \int_{j\epsilon}^{(j+1)\epsilon} dt \int d^3x\, H(\vec{x}, t). \tag{B.18}$$

We see that the formula is analogous to the identity above satisfied by the exponencial, and we may write

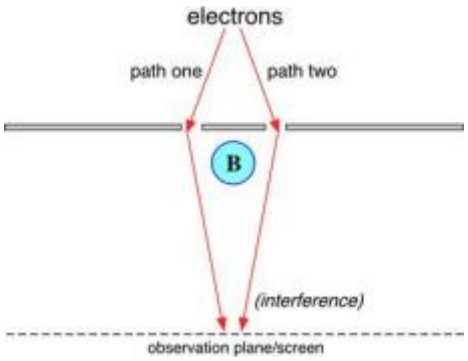
$$S = \mathcal{T}\exp\left(\sum_{j=-\infty}^{j=\infty}\right) = \mathcal{T}\exp\left(-\frac{i}{\hbar}\int dx\, H(x)\right). \tag{B.19}$$

The only subtlety we had to include was the time-ordering operator $\mathcal{T}$ because the factors in the product definig S above were time-ordered, too (and operators do no commute in general) and the operator $\mathcal{T}$ guarantees that this ordering will be preserved.

Appendix C

The Bohm-Aharonov effect and the topology of the vacuum

The famous two-slit experiment with electrons shows the wave nature of the elctrons producing a characteristic interference pattern. The idea of Bohm and Aharonov was to introduce a small solenoid behind the wall between the slits.



There are lines of magnetic field **B** inside the solenoid, but not outside, so, as long as the solenoid is small enough, the electrons always move on a field-free region. But the presence of the solenoid causes a shift in the interference pattern, even though the electrons only ever move through regions of no magnetic field. The significance of this effect is that, in quantum theory, an electron is influed by the vector potential **A**, even though it travels entirely regions where $\mathbf{B} = 0$. The Bohm-Aharonov effect owes its existence to the non-trivial topology of the vacuum, and the fact that electrodynamics is a gauge theory.

Outside the solenoid, $\mathbf{E} = 0$ and $\mathbf{B} = 0$ so the energy density of the electromagnetic field $U = 0$ and we have a vacuum. On the other hand, $\mathbf{A} \neq 0$ so the vacuum has a structure "structure". Since $\nabla \times \mathbf{A} = 0$, we may write $\mathbf{A} = \nabla \times f$ for some function f which is given by

$$f = \frac{BR^2}{2}\phi + c \tag{9.4}$$

where c is a constant of integration. f is not a single-valued function, since it increases by $\pi R^2 B$ when $\phi \to \phi + 2\pi$.

Regular non-single-valued functions, however, may only exist in non-simply connected spaces. A simple connected space is one in which all closed curves may be shrunk continously to a point. A non-simple connected space is one which not all curves may be continously shrunk to a point. The relevant space in this problem is the space of the vacuum, i.e. the space outside the solenoid, and that is not simply connected. The mathematical reason for this is a that the configuration space of the null field (vacuum) is the plane with a hole in: which is $S^1 \times \mathbb{R}$.

It is thus an essential condition for the Bohm-Aharonov effect to occur that the configuration space of the vacuum is not simply connected. The group space of $U(1)$ is the circle, which is denote by S^1. This group space is not simply connected, because a path which goes twice round a circle cannot be continously deformed into one which goes round once. In other words, it is becuse the gauge group of electromagnetism, U_1, is not simply connected that the Bohm-Aharonov effect is possible. The mathematical reason for this is a that the configuration space of the null field (vacuum) is the plane with a hole in: which is $S^1 \times \mathbb{R}$.

Bibliography

[1] Óscar García-Prada, *EXISTENCIA DE YANG-MILLS Y DEL SALTO DE MASA*. Jornadas sobre los problemas del milenio (Barcelona, 2011).

[2] Lewis H. Ryder, *QUANTUM FIELD THEORY*. ISBN 0-521-23764-5 Cambridge University Press (1985).

[3] Stefan Pokorski, *Gauge Field Theories*. ISBN 0-521-26537-1 Cambridge University Press (1987).

[4] Arthur Jaffe, Edward Witten, *Quantum Yang-Mills theory*, in *THE MILLENIUM PRIZE PROBLEMS* (AMS, 2006), 129-152.

[5] A. Pich, *The Standard Model of Electroweak Interactions*.Arxiv preprint arXiv:0705.4264 (2007).

[6] David Bleecker, *GAUGE THEORY and VARIATIONAL PRINCIPLES*. ISBN 0-486-44546-1 Dover Publications, inc. (1981)

[7] Arthur Jaffe, *Quantum Theory and Relativity*. Cambridge, MA 02138, USA (2007).

[8] Arthur Jaffe, *Constructive Quantum Field Theory*. Harvard University, Cambridge, MA 02138, USA.

[9] Lochlaim O'Raifeartaigh, *The Dawning of Gauge Theory*. ISBN 0-691-02978-4 Princeton University Press (1997).

[10] C. W. Misner & K. Thorne and J.A Wheeler, *GRAVITATION*. W. H. FREEMAN AND COMPANY (San Fracisco).

[11] F.Schwabl, *Advanced Quantum Mechanics*. ISBN 3-540-64478-4 Springer-Verlag Berlin Heidelberg New York (1999).

[12] Frank W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*. ISBN 0-387-90894-3 Springer-Verlag New York Berlin Heidelberg (1983).

[13] Daniel S. Freed & Karen K. Uhlenbeck (Editors),*Geometry and Quantum Field Theory*. ISBN 0-8218-0400-6 IAS/PARK CITY

[14] Eric D'Hoker and D.H. Phong, *Lectures on Supersymmetric Yangs-Mills Theory and Integrable Systems*. UCLA/99/TEP/28 Columbia/Math/99

[15] M. Göckeler & T. Schücker, *Differential geometry, gauge theories, and gravity*. ISBN 0-521-32960-4 Cambridge University Press (1987)

[16] Ta-Pei Cheng & Ling-Fong Li, Gauge theory of elementary particle physics. ISBN 0-19-851956-7 Oxford University Press (1984)